

(12)特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関  
国際事務局(43) 国際公開日  
2003年11月6日 (06.11.2003)

PCT

(10) 国際公開番号  
WO 03/091989 A1(51) 国際特許分類<sup>7</sup>:

G10L 19/00

(74) 代理人: 鷲田 公一 (WASHIDA,Kimihito); 〒206-0034  
東京都 多摩市 鶴牧1丁目24-1 新都市センタービル  
5階 Tokyo (JP).

(21) 国際出願番号:

PCT/JP03/05419

(22) 国際出願日:

2003年4月28日 (28.04.2003)

(25) 国際出願の言語:

日本語

(81) 指定国 (国内): AE, AG, AL, AM, AT, AU, AZ, BA, BB,  
BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK,  
DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU,  
ID, IL, IN, IS, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT,  
LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO,  
NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL,  
TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU,  
ZA, ZM, ZW.

(30) 優先権データ:

特願2002-127541 2002年4月26日 (26.04.2002) JP

特願2002-267436 2002年9月12日 (12.09.2002) JP

(84) 指定国 (広域): ARIPO 特許 (GH, GM, KE, LS, MW, MZ,  
SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア特許 (AM,  
AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ特許  
(AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB,  
GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR),  
OAPI 特許 (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW,  
ML, MR, NE, SN, TD, TG).(71) 出願人 (米国を除く全ての指定国について): 松下電  
器産業株式会社 (MATSUSHITA ELECTRIC INDUS-  
TRIAL CO., LTD.) [JP/JP]; 〒571-8501 大阪府門真市  
大字門真1006番地 Osaka (JP).

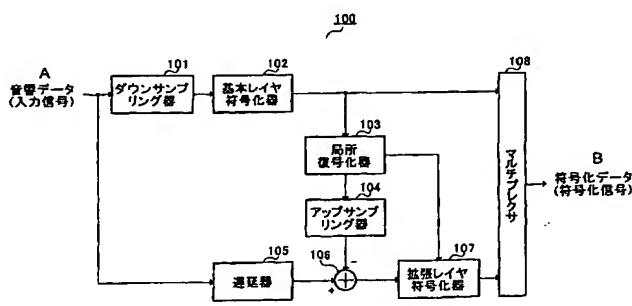
(72) 発明者; および

(75) 発明者/出願人 (米国についてのみ): 押切 正浩 (OS-  
HIKIRI,Masahiro) [JP/JP]; 〒239-0841 神奈川県横須  
賀市野比1-31-7-A Kanagawa (JP).

[統葉有]

(54) Title: CODING DEVICE, DECODING DEVICE, CODING METHOD, AND DECODING METHOD

(54) 発明の名称: 符号化装置、復号化装置及び符号化方法、復号化方法



A...ACOUSTIC DATA (INPUT SIGNAL)  
101...DOWN-SAMPLING DEVICE  
102...BASIC LAYER CODING DEVICE  
103...LOCAL DECODING DEVICE  
104...UP-SAMPLING DEVICE  
105...DELAY DEVICE  
107...EXTENDED LAYER CODING DEVICE  
108...MULTIPLEXER  
B...ENCODED DATA (ENCODED SIGNAL)

WO 03/091989 A1

(57) Abstract: A down-sampling device (101) down-samples the input signal sampling rate FH to a sampling rate FL. A basic layer coding device (102) encodes an acoustic signal of the sampling rate FL. A local decoding device (103) decodes an encoded code output from the basic layer coding device (102). An up-sampling device (104) increases the sampling rate of the decoded signal to FH. A subtractor (106) subtracts the decoded signal from the acoustic signal of the sampling rate FH. An extended layer coding device (107) encodes the signal output from the subtractor (106) by using the decoding result parameter output from the local decoding device (103).

[統葉有]



添付公開書類:  
— 國際調査報告書

2文字コード及び他の略語については、定期発行される各PCTガゼットの巻頭に掲載されている「コードと略語のガイダンスノート」を参照。

---

(57) 要約:

ダウンサンプリング器101は、入力信号のサンプリングレートをサンプリングレートFHからサンプリングレートFLにダウンサンプリングする。基本レイヤ符号化器102は、サンプリングレートFLの音響信号を符号化する。局所復号化器103は、基本レイヤ符号化器102から出力された符号化コードを復号する。アップサンプリング器104は、復号信号のサンプリングレートをFHに上げる。減算器106は、サンプリングレートFHの音響信号から復号信号を減算する。拡張レイヤ符号化器107は、減算器106から出力された信号を局所復号化器103から出力された復号結果のパラメータを用いて符号化する。

## 明細書

## 符号化装置、復号化装置及び符号化方法、復号化方法

## 5 技術分野

本発明は、楽音信号または音声信号などの音響信号を高能率に圧縮符号化及び復号化を行う符号化装置、復号化装置及び符号化方法、復号化方法に関し、特に符号化コードの一部からでも楽音や音声を復号することができるスケーラブル符号化及び復号化に好適な符号化装置、復号化装置及び符号化方法、復号化方法に関するもの。

## 背景技術

楽音信号または音声信号を低ビットレートで圧縮する音響符号化技術は、移動体通信における電波等の伝送路容量及び記録媒体の有効利用のために重要である。音声信号を符号化する音声符号化に、ITU(International Telecommunication Union)で規格化されている G726、G729 などの方式が存在する。これらの方は、狭帯域信号(300Hz～3.4kHz)を対象とし、8kbit/s～32kbit/s のビットレートで高品質に符号化ができる。

また、広帯域信号(50Hz～7kHz)を対象とする標準方式として ITU の G722、G722.1 や、3GPP(The 3rd Generation Partnership Project)の AMR-WB などが存在する。これらの方は、ビットレートが 6.6kbit/s～64kbit/s で広帯域音声信号を高品質に符号化ができる。

ここで、音声信号を低ビットレートで高能率に符号化を行う有効な方法に、C E L P (Code Excited Linear Prediction)がある。C E L P は、人間の音声生成モデルを工学的に模擬したモデルに基づき符号化を行う方法である。具体的には、C E L P は、乱数で表される励振信号を周期性の強さに対応するピッチ

フィルタと声道特性に対応する合成フィルタに通し、その出力信号と入力信号の二乗誤差が聴覚特性の重み付けの下で最小になるよう符号化パラメータを決定する。

そして、最近の標準音声符号化方式の多くが C E L P に基づいた符号化を行っている。例えは、G729 は、8kbit/s で狭帯域信号の符号化を行うことができ、  
5 AMR-WB は 6.6kbit/s～23.85kbit/s で広帯域信号を符号化を行うことができる。

一方で、楽音信号を符号化する楽音符号化の場合は、M P E G (Moving Picture Expert Group)で規格化されているレイヤ III 方式や A A C 方式のように、楽音信号を周波数領域に変換し、聴覚心理モデルを利用して符号化を行う  
10 方法が一般的である。これら方式は、サンプリングレートが 44.1kHz の信号に対しチャネル当たり 64kbit/s～96kbit/s でほとんど劣化が生じないことが知られている。

この楽音符号化は、音楽に対して高品質に符号化を行う方式である。楽音符号化は、前述の説明にある背景に音楽や環境音がある音声信号についても高品質に符号化を行うことができる。そして、対象となる信号の帯域も C D 品質である 22kHz 程度まで対応可能である。  
15

しかしながら、音声信号が主体で、背景に音楽や環境音が重畠している信号に音声符号化方式を用いて符号化する場合、背景部の音楽や環境音の影響で、  
20 背景部の信号のみならず音声信号も劣化してしまい全体的な品質が低下するという問題がある。

この問題は、音声符号化方式が、C E L P という音声モデルに特化した方式を基本にしているために生じる。また、音声符号化方式が対応できる信号帯域は 7kHz までであり、それ以上の帯域の成分を持つ信号に対しては構成上十分に対応しきれないという問題があった。

25 また、楽音符号化方式では、高品質な符号化を実現するためにはビットレートを高くして使用する必要がある。楽音符号化方式では、仮にビットレートを

32kbit/s 程度まで低く抑えて符号化した場合、復号信号の品質が大きく低下するという問題がある。そのため、伝送レートの低い通信網で使用できないという問題がある。

## 5 発明の開示

本発明の目的は、音声が主体で背景に音楽や環境音が重畠しているような信号であっても低ビットレートで高品質に符号化及び復号化できる符号化装置、復号化装置及び符号化方法、復号化方法を提供することである。

この目的は、基本レイヤと拡張レイヤの 2 つの階層を持ち、基本レイヤで C E L P を基にして入力信号の狭帯域または広帯域の周波数領域を低いビットレートで高品質に符号化し、基本レイヤで表しきれない背景の音楽や環境音、さらに基本レイヤでカバーする周波数領域よりも高い周波数成分の信号を拡張レイヤで符号化を行うことにより達成される。

## 15 図面の簡単な説明

図 1 は、本発明の実施の形態 1 に係る信号処理装置の構成を示すブロック図、

図 2 は、入力信号の成分の一例を示す図、

図 3 は、上記実施の形態に係る信号処理装置の信号処理方法の一例を示す図、

図 4 は、基本レイヤ符号化器の構成の一例を示す図、

図 5 は、拡張レイヤ符号化器の構成の一例を示す図、

図 6 は、拡張レイヤ符号化器の構成の一例を示す図、

図 7 は、拡張 L P C 係数算出の一例を示す図、

図 8 は、本発明の実施の形態 3 に係る信号処理装置の拡張レイヤ符号化器の構成を示すブロック図、

図 9 は、本発明の実施の形態 4 に係る信号処理装置の拡張レイヤ符号化器の構成を示すブロック図、

図 10 は、本発明の実施の形態 5 に係る信号処理装置の構成を示すブロック図、

図 11 は、基本レイヤ復号化器の一例を示すブロック図、

図 12 は、拡張レイヤ復号化器の一例を示すブロック図、

5 図 13 は、拡張レイヤ復号化器の構成の一例を示す図、

図 14 は、本発明の実施の形態 7 に係る信号処理装置の拡張レイヤ復号化器の構成を示すブロック図、

図 15 は、本発明の実施の形態 8 に係る信号処理装置の拡張レイヤ復号化器の構成を示すブロック図、

10 図 16 は、本発明の実施の形態 9 に係る音響符号化装置の構成を示すブロック図、

図 17 は、音響信号の情報の分布の一例を示す図、

図 18 は、基本レイヤと拡張レイヤで符号化の対象とする領域の一例を示す図、

15 図 19 は、音響（音楽）信号のスペクトルの一例を示す図、

図 20 は、上記実施の形態の音響符号化装置の周波数決定部の内部構成の一例を示すブロック図、

図 21 は、上記実施の形態の音響符号化装置の聴覚マスキング算出器の内部構成の一例を示す図、

20 図 22 は、上記実施の形態の拡張レイヤ符号化器の内部構成の一例を示すブロック図、

図 23 は、上記実施の形態の聴覚マスキング算出器の内部構成の一例を示すブロック図、

25 図 24 は、本発明の実施の形態 9 に係る音響復号化装置の構成を示すブロック図、

図 25 は、上記実施の形態の音響復号化装置の拡張レイヤ復号化器の内部構

成の一例を示すブロック図、

図 2 6 は、本発明の実施の形態 1 0 の基本レイヤ符号化器の内部構成の一例を示すブロック図、

図 2 7 は、上記実施の形態の基本レイヤ復号化器の内部構成の一例を示すブロック図、

図 2 8 は、上記実施の形態の基本レイヤ復号化器の内部構成の一例を示すブロック図、

図 2 9 は、本発明の実施の形態 1 1 に係る音響符号化装置の周波数決定部の内部構成の一例を示すブロック図、

図 3 0 は、上記実施の形態の推定誤差スペクトル算出器が算出する残差スペクトルの一例を示す図、

図 3 1 は、本発明の実施の形態 1 2 に係る音響符号化装置の周波数決定部の内部構成の一例を示すブロック図、

図 3 2 は、上記実施の形態の音響符号化装置の周波数決定部の内部構成の一例を示すブロック図、

図 3 3 は、本発明の実施の形態 1 3 に係る音響符号化装置の拡張レイヤ符号化器の内部構成の一例を示すブロック図、

図 3 4 は、上記実施の形態の順序づけ部の推定歪値の順位づけの一例を示す図、

図 3 5 は、本発明の実施の形態 1 3 に係る音響復号化装置の拡張レイヤ復号化器の内部構成の一例を示すブロック図、

図 3 6 は、本発明の実施の形態 1 4 に係る音響符号化装置の拡張レイヤ符号化器の内部構成の一例を示すブロック図、

図 3 7 は、本発明の実施の形態 1 4 に係る音響復号化装置の拡張レイヤ復号化器の内部構成の一例を示すブロック図、

図 3 8 は、上記実施の形態の音響符号化装置の周波数決定部の内部構成の一

例を示すブロック図、

図39は、本発明の実施の形態14に係る音響復号化装置の拡張レイヤ復号化器の内部構成の一例を示すブロック図、

図40は、本発明の実施の形態15に係る通信装置の構成を示すブロック図、

5 図41は、本発明の実施の形態16に係る通信装置の構成を示すブロック図、

図42は、本発明の実施の形態17に係る通信装置の構成を示すブロック図、

及び、

図43は、本発明の実施の形態18に係る通信装置の構成を示すブロック図である。

10

#### 発明を実施するための最良の形態

本発明の骨子は、基本レイヤと拡張レイヤの2つの階層を持ち、基本レイヤはC E L Pを基にして入力信号の狭帯域または広帯域の周波数領域を低いビットレートで高品質に符号化する。次に、基本レイヤで表しきれない背景の音楽や環境音、さらに基本レイヤでカバーする周波数領域よりも高い周波数成分の信号を拡張レイヤにおいて符号化を行い、拡張レイヤは楽音符号化方式のようにあらゆる信号に対応できる構成にすることである。

これにより、基本レイヤで表現しきれない背景の音楽や環境音、さらに基本レイヤでカバーする周波数領域よりも高い周波数成分の信号を効率よく符号化を行うことができる。この際、基本レイヤの符号化コードより得られる情報を利用して拡張レイヤの符号化を行うことが本発明の特徴である。これにより、拡張レイヤの符号化ビット数を低く抑えることができるという効果が得られる。

以下、本発明の実施の形態について、図面を用いて説明する。

#### (実施の形態1)

25 図1は、本発明の実施の形態1に係る信号処理装置の構成を示すブロック図である。図1の信号処理装置100は、ダウンサンプリング器101と、基本

レイヤ符号化器 102 と、局所復号化器 103 と、アップサンプリング器 104 と、遅延器 105 と、減算器 106 と、拡張レイヤ符号化器 107 と、マルチプレクサ 108 とから主に構成される。

5 ダウンサンプリング器 101 は、入力信号のサンプリングレートをサンプリングレート  $F_H$  からサンプリングレート  $F_L$  にダウンサンプリングし、サンプリングレート  $F_L$  の音響信号を基本レイヤ符号化器 102 に出力する。ここで、サンプリングレート  $F_L$  は、サンプリングレート  $F_H$  より低い周波数である。

基本レイヤ符号化器 102 は、サンプリングレート  $F_L$  の音響信号を符号化し、符号化コードを局所復号化器 103 とマルチプレクサ 108 に出力する。

10 局所復号化器 103 は、基本レイヤ符号化器 102 から出力された符号化コードを復号し、復号信号をアップサンプリング器 104 に出力し、復号した結果求められるパラメータを拡張レイヤ符号化器 107 に出力する。

アップサンプリング器 104 は、復号信号のサンプリングレートを  $F_H$  に上げて減算器 106 に出力する。

15 遅延器 105 は、入力されたサンプリングレート  $F_H$  の音響信号を所定の時間の遅延した後、減算器 106 する。この遅延時間をダウンサンプリング器 101 と基本レイヤ符号化器 102 と局所復号化器 103 とアップサンプリング器 104 で生じる時間遅れと同値とすることにより、次の減算処理での位相のずれを防ぐ。

20 減算器 106 は、サンプリングレート  $F_H$  の音響信号から復号信号を減算し、減算結果を拡張レイヤ符号化器 107 に出力する。

拡張レイヤ符号化器 107 は、減算器 106 から出力された信号を局所復号化器 103 から出力された復号結果のパラメータを用いて符号化し、マルチプレクサ 108 に出力する。マルチプレクサ 108 は、基本レイヤ符号化器 102 と拡張レイヤ符号化器 107 とにおいて符号化された信号を多重化して出力する。

次に、基本レイヤ符号化と拡張レイヤ符号化について説明する。図2は、入力信号の成分の一例を示す図である。図2において、縦軸は、信号の成分の情報量を示し、横軸は、周波数を示す。図2では、入力信号に含まれる音声情報と、背景音楽/背景雑音情報がどの周波数帯に存在しているかを表している。

5 音声情報は、周波数の低い領域に情報が多く存在し、高域に向かうほど情報量は減少する。それに対し、背景音楽・背景雑音情報は、音声情報と比べると相対的に低域の情報は少なく、高域に含まれる情報が大きい。

そこで、本発明の信号処理装置は、複数の符号化方式を用い、それぞれの符号化方式が適する領域毎に異なる符号化を行う。

10 図3は、本実施の形態に係る信号処理装置の信号処理方法の一例を示す図である。図3において、縦軸は、信号の成分の情報量を示し、横軸は、周波数を示す。

15 基本レイヤ符号化器102は、0～FL間の周波数帯の音声情報を効率よく表すように設計されており、この領域での音声情報は品質良く符号化することができる。しかし、0～FL間の周波数帯の背景音楽・背景雑音情報の符号化品質は高くない。拡張レイヤ符号化器107は、基本レイヤ符号化器102で符号化できない部分と、FH～FL間の周波数帯の信号を符号化する。

20 よって、基本レイヤ符号化器102と拡張レイヤ符号化器107とを組み合わせることで広い帯域で高品質な符号化が実現できる。さらに、少なくとも基本レイヤ符号化手段の符号化コードだけでも音声情報が復号できるというスケーラブルな機能が実現できる。

25 このように、局所復号化器103における符号化で生成されたパラメータのうち有用なものを拡張レイヤ符号化器107に与え、拡張レイヤ符号化器107は、このパラメータを利用して符号化を行う。

このパラメータは符号化コードから生成されるため、本実施の形態の信号処理装置により符号化された信号を復号する場合に、音響復号化の過程で同じパ

ラメータを得ることができ、このパラメータを付加して復号側に伝送する必要がない。このために、拡張レイヤ符号化手段は付加情報の増加を伴うことなしに、符号化処理の効率化を図ることができる。

例えば、局所復号化器 103において復号されるパラメータのうち、拡張レイヤ符号化器 107で用いられるパラメータとして、入力信号が母音のように周期性の強い信号か子音のように雑音性の強い信号かを表す有声／無声フラグを使う構成がある。有声／無声フラグを用い、有声である区間において拡張レイヤでは高域よりも低域を重視してビット配分を行い、無声である区間では低域よりも高域を重視してビット配分を行う、などの適応化を図ることができる。

10 このように、本実施の形態の信号処理装置によれば、入力信号から所定の周波数以下の成分を取り出して音声符号化に適した符号化を行い、得られた符号化コードを復号した結果を用いて楽音符号化に適した符号化を行うことにより、低ビットレートで高品質に符号化を行うことができる。

また、サンプリングレート  $F_H$  と  $F_L$  には、 $F_H$  が  $F_L$  より大きい値であれば良く、値は限定されない。例えば、サンプリングレートを  $F_H=24\text{kHz}$ 、 $F_L=16\text{kHz}$  とし、符号化することができる。

#### (実施の形態 2)

本実施の形態では、実施の形態 1 の局所復号化器 103で復号されるパラメータのうち、拡張レイヤ符号化器 107で用いられるパラメータとして、入力信号のスペクトルを表す LPC 係数を用いる例について説明する。

本実施の形態の信号処理装置は、図 1 の基本レイヤ符号化器 102において CELP を用いた符号化を行い、拡張レイヤ符号化器 107において入力信号のスペクトルを表す LPC 係数を用いて符号化を行う。

ここでは最初に、基本レイヤ符号化器 102 の詳細な動作の説明を行った後に、拡張レイヤ符号化器 107 の基本構成の説明を行う。ここでいう基本構成とは、今後の実施の形態の説明を簡便にするためのもので、局所復号化器 10

3の符号化パラメータを用いない構成を指す。その後に本実施の形態の特徴である局所復号化器103でLPC係数を復号し、このLPC係数を用いた拡張レイヤ符号化器107の説明を行う。

図4は、基本レイヤ符号化器102の構成の一例を示す図である。図4の基  
5 本レイヤ符号化器102は、LPC分析器401と、聴感重み付け部402と、  
適応符号帳探索器403と、適応ゲイン量子化器404と、目標ベクトル生成  
器405と、雑音符号帳探索器406と、雑音ゲイン量子化器407と、マル  
チプレクサ408とから主に構成される。

10 LPC分析器401は、ダウンサンプリング器101においてサンプリング  
レートFLでサンプリングされた入力信号からLPC係数を求め、聴感重み付  
け部402に出力する。

15 聴感重み付け部402は、LPC分析器401で求められたLPC係数を基  
に入力信号に重み付けを行い、重み付けされた入力信号を適応符号帳探索器4  
03、適応ゲイン量子化器404、及び目標ベクトル生成器405に出力する。  
適応符号帳探索器403は、聴覚重み付けされた入力信号を目標信号として  
20 適応符号帳の探索を行い、探索した適応ベクトルを適応ゲイン量子化器404  
と目標ベクトル生成器405に出力する。そして、適応符号帳探索器403は、  
量子化歪みが最も少ないとされた適応ベクトルのコードをマルチプレクサ40  
8に出力する。

25 適応ゲイン量子化器404は、適応符号帳探索器403から出力される適応  
ベクトルに乘算する適応ゲインを量子化し、目標ベクトル生成器405に出力  
する。そして、そのコードをマルチプレクサ408に出力する。

目標ベクトル生成器405は、聴感重み付け部402から出力された入力信  
号を適応ベクトルに適応ゲインを乗算した結果でベクトル減算し、減算結果を  
25 目標ベクトルとして雑音符号帳探索器406と雑音ゲイン量子化器407に出  
力する。

雑音符号帳探索器 406 は、目標ベクトル生成器 405 から出力された目標ベクトルとの歪が最小となる雑音ベクトルを雑音符号帳の中から探索する。そして、雑音符号帳探索器 406 は、探索した雑音ベクトルを雑音ゲイン量子化器 407 に与えるとともに、そのコードをマルチプレクサ 408 に出力する。

5 雜音ゲイン量子化器 407 は、雑音符号帳探索器 406 において探索された雑音ベクトルに乘算する雑音ゲインを量子化し、そのコードをマルチプレクサ 408 に出力する。

マルチプレクサ 408 は、LPC 係数、適応ベクトル、適応ゲイン、雑音ベクトル、雑音ゲインの符号化コードを多重化して局所復号化器 103 及びマルチプレクサ 108 に出力する。

10 次に、図 4 の基本レイヤ符号化器 102 の動作について説明する。最初に、ダウンサンプリング器 101 から出力されたサンプリングレート  $F_L$  の信号が入力され、LPC 分析器 401 において LPC 係数が求められる。この LPC 係数は、LSP 係数などの量子化に適したパラメータに変換され、量子化される。この量子化で得られる符号化コードがマルチプレクサ 408 に与えられ、かつ符号化コードから量子化後の LSP 係数が算出され LPC 係数に変換される。

15 この変換により、量子化後の LPC 係数が求められる。この量子化後の LPC 係数を利用して、適応符号帳、適応ゲイン、雑音符号帳および雑音ゲインの符号化を行う。

20 次に、聴感重み付け部 402 は、LPC 分析器 401 で求められた LPC 係数に基づいて入力信号に重み付けを行う。この重み付けは、量子化歪のスペクトルを入力信号のスペクトル包絡にマスクされるようスペクトル整形を行うことを目的として成される。

25 次に、聴覚重み付けされた入力信号を目標信号とし、適応符号帳探索器 403 において適応符号帳が探索される。過去の音源系列をピッチ周期で繰り返し

た信号を適応ベクトルと呼び、あらかじめ定められた範囲のピッチ周期で生成された適応ベクトルによって適応符号帳は構成される。

聴覚重み付けされた入力信号を  $t(n)$ 、ピッチ周期  $i$  の適応ベクトルに L P C 係数で構成される重み付き合成フィルタのインパルス応答を畳み込んだ信号を 5  $pi(n)$ としたとき、次の式（1）の評価関数  $D$  を最小とする適応ベクトルのピッ チ周期  $i$  がパラメータとしてマルチプレクサ 408 に送られる。

$$D = \sum_{n=0}^{N-1} t^2(n) - \frac{\left( \sum_{n=0}^{N-1} t(n) pi(n) \right)^2}{\sum_{n=0}^{N-1} pi^2(n)} \quad \cdots (1)$$

ここで、 $N$  はベクトル長を示す。

10 次に、適応ベクトルに乘じられる適応ゲインの量子化を適応ゲイン量子化器 404において行う。適応ゲイン  $\beta$  は次の式（2）で表され、この  $\beta$  をスカラ 一量子化して、その符号がマルチプレクサ 408 に送られる。

$$\beta = \frac{\sum_{n=0}^{N-1} t(n) pi(n)}{\sum_{n=0}^{N-1} pi^2(n)} \quad \cdots (2)$$

次に、目標ベクトル生成器 405において入力信号から適応ベクトルの影響を減算して、雑音符号帳探索器 406と雑音ゲイン量子化器 407で用いる目標ベクトルを生成する。ここで、 $pi(n)$ を式1で表される評価関数Dを最小とするときの適応ベクトルに合成フィルタを畳み込んだ信号、 $\beta q$ を式2で表される適応ベクトル $\beta$ をスカラー量子化したときの量子化値としたとき、目標ベクトル $t2(n)$ は次の式(3)に表される。

$$t2(n) = t(n) - \beta q \cdot pi(n) \quad \dots (3)$$

10

前記目標ベクトル $t2(n)$ とLPC係数が、雑音符号帳探索器 406に与えられ、雑音符号帳の探索が行われる。

ここで、雑音符号帳探索器 406が備える雑音符号帳の代表的な構成に代数(Algebraic)符号帳がある。代数符号帳とは、振幅1のパルスをあらかじめ定められた非常に少ない数だけ有するベクトルで表される。さらに、代数符号帳は、パルスごとにとりうる位置は重複することなくあらかじめ決められている。そして、代数符号帳は、パルスの位置とパルスの符号(極性)の最適な組み合わせを少ない計算量で決定することができるという特徴がある。

目標ベクトルを $t2(n)$ 、コード $j$ に対応する雑音ベクトルに重み付き合成フィルタのインパルス応答を畳み込んだ信号を $cj(n)$ としたとき、次の式(4)の評価関数Dを最小とする雑音ベクトルのインデックス $j$ がパラメータとしてマルチプレクサ 408に送られる。

$$D = \sum_{n=0}^{N-1} t_2^2(n) - \frac{\left( \sum_{n=0}^{N-1} t_2(n) c_j(n) \right)^2}{\sum_{n=0}^{N-1} c_j^2(n)} \quad \dots (4)$$

次に、雑音ベクトルに乘じられる雑音ゲインの量子化を雑音ゲイン量子化器  
 407において行う。雑音ゲイン  $\gamma$  は次の式 (5) で表され、この  $\gamma$  をスカラ  
 5 一量子化して、その符号がマルチプレクサ 408 に送られる。

$$\gamma = \frac{\sum_{n=0}^{N-1} t_2(n) c_j(n)}{\sum_{n=0}^{N-1} c_j^2(n)} \quad \dots (5)$$

マルチプレクサ 408 は、送られてきた L P C 係数、適応符号帳、適応ゲイ  
 10 ン、雑音符号帳、雑音ゲインの符号化コードを多重化して局所復号化器 103  
 及びマルチプレクサ 108 に出力する。

そして、新しい入力信号が存在する間、上記処理を繰り返す。新しい入力信  
 号が存在しない場合には、処理を終了する。

次に、拡張レイヤ符号化器 107 について説明する。図 5 は、拡張レイヤ符  
 15 号化器 107 の構成の一例を示す図である。図 5 の拡張レイヤ符号化器 107

は、LPC分析器501と、スペクトル包絡算出器502と、MDCT部503と、パワー算出器504と、パワー正規化器505と、スペクトル正規化器506と、Barkスケール正規化器508と、Barkスケール形状算出器507と、ベクトル量子化器509と、マルチプレクサ510とから主に構成される。

5 LPC分析器501は、入力信号にLPC分析を行い、得られたLPC分析係数をスペクトル包絡算出器502及びマルチプレクサ510に出力する。スペクトル包絡算出器502は、LPC係数からスペクトル包絡を算出してベクトル量子化器509に出力する。

10 MDCT部503は、入力信号にMDCT変換 (Modified Discrete Cosine Transform: 変形離散コサイン変換) を行い、得られたMDCT係数をパワー算出器504及びパワー正規化器505に出力する。パワー算出器504は、MDCT係数のパワーを求め、量子化した後、パワー正規化器505及びマルチプレクサ510に出力する。

15 パワー正規化器505は、量子化後のパワーにてMDCT係数を正規化し、正規化後のパワーをスペクトル正規化器506に出力する。スペクトル正規化器506は、スペクトル包絡を用いてパワーにより正規化されたMDCT係数を正規化し、Barkスケール形状算出器507及びBarkスケール正規化器508に出力する。

20 Barkスケール形状算出器507は、Barkスケールにて等間隔に帯域分割されたスペクトルの形状を算出した後に、前記スペクトル形状を量子化し、量子化したスペクトル形状をBarkスケール正規化器508、ベクトル量子化器509、及びマルチプレクサ510に出力する。

25 Barkスケール正規化器508は、各帯域のBarkスケール形状B(k)を量子化し、その符号化コードをマルチプレクサ510に出力する。そして、Barkスケール正規化器508は、Barkスケール形状を復号化して正規化MDCT係数を生成し、ベクトル量子化器509に出力する。

ベクトル量子化器 509 は、Bark スケール正規化器 508 から出力された正規化MDCT係数をベクトル量子化し、最も歪が小さい代表値を求め、このインデックスを符号化コードとしてマルチプレクサ 510 に出力する。

マルチプレクサ 510 は、符号化コードを多重化して、マルチプレクサ 10 5 8 に出力する。

次に、図 5 の拡張レイヤ符号化器 107 の動作について説明する。図 1 の減算器 106 で得られる減算信号が、LPC 分析器 501 において LPC 分析される。そして、LPC 分析により LPC 係数が算出される。この LPC 係数を LSP 係数などの量子化に適したパラメータに変換した後に量子化を行う。こので得られた LPC 係数に関する符号化コードはマルチプレクサ 510 に与えられる。

スペクトル包絡算出器 502 では、復号された LPC 係数を基に、次の式(6)に従いスペクトル包絡を算出する。

$$15 \quad env(m) = \left| \frac{1}{1 - \sum_{i=1}^{NP} \alpha_q(i) e^{-j \frac{2\pi m i}{M}}} \right| \quad \cdots (6)$$

ここで  $\alpha_q$  は、復号された LPC 係数をしめし、NP は LPC 係数の次数、M はスペクトル分解能を示す。式(6)により得られたスペクトル包絡 env(m) は、後に説明するスペクトル正規化器 506 およびベクトル量子化器 509 で利用

される。

次に、入力信号は、MDCT部503においてMDCT変換が行われ、MDCT係数が求められる。MDCT変換は、前後の隣接フレームと分析フレームを半分ずつ完全に重ね合わせ、かつ分析フレームの前半部は奇関数、後半部は偶関数という直交基底を用いるため、フレーム境界歪が発生しないという特徴 5 がある。MDCTを行う際には、sin窓などの窓関数を入力信号に乘ずる。MDCT係数を  $X(m)$  とすると、MDCT係数は次の式 (7) に従い算出される。

$$X(m) = \sqrt{\frac{1}{N}} \sum_{n=0}^{2N-1} x(n) \cos \left\{ \frac{(2n+1+N) \cdot (2m+1)\pi}{4N} \right\} \quad \dots (7)$$

10

ここで  $x(n)$  は、入力信号に窓関数を乗算した信号を示す。

次に、パワー算出器504では、MDCT係数  $X(m)$  のパワーを求め量子化する。そして、パワー正規化器505が、式 (8) を用い、当該量子化後のパワーにてMDCT係数を正規化する。

15

$$pow = \sum_{m=0}^{M-1} X(m)^2 \quad \dots (8)$$

ここで、 $M$  はMDCT係数の次数を示す。MDCT係数のパワー  $pow$  を量子化

した後に、この符号化コードをマルチプレクサ 510 に送る。符号化コードを使ってMDCT係数のパワーを復号した後に、その値を用いてMDCT係数を次の式 (9) に従い正規化する。

5      
$$X1(m) = \frac{X(m)}{\sqrt{powq}} \quad \dots (9)$$

ここで、 $X1(m)$  はパワー正規化後のMDCT係数を表し、 $powq$  は量子化後のMDCT係数のパワーを示す。

次に、スペクトル正規化器 506 は、スペクトル包絡を用いてパワーにより 10 正規化されたMDCT係数を正規化する。スペクトル正規化器 506 では次の式 (10) に従い正規化を行う。

$$X2(m) = \frac{X1(m)}{env(m)} \quad \dots (10)$$

15      次に、Bark スケール形状算出器 507 は、Bark スケールにて等間隔に帯域分割されたスペクトルの形状を算出した後に、前記スペクトル形状を量子化する。 Bark スケール形状算出器 507 は、この符号化コードをマルチプレクサ 510 に送ると共にその復号値を用いてスペクトル正規化器 506 の出力信号である MDCT 係数  $X2(m)$  の正規化を行う。Bark スケールと Herz スケールは次の式 (1 20 1) で表される変換式にて対応付けされる。

$$B = 13 \tan^{-1}(0.76f) + 3.5 \tan^{-1}\left(\frac{f}{7.5}\right) \quad \dots (11)$$

ここで B は Bark スケールを、 f は Herz スケールを示す。 Bark スケール形状算出器 507 は、 Bark スケールで等間隔に帯域分割されたサブバンドそれぞれに

5 対し、次の式 (12) に従い形状を算出する。

$$B(k) = \sum_{m=f_l(k)}^{f_h(k)} X2(m)^2 \quad 0 \leq k < K \quad \dots (12)$$

10 ここで  $f_l(k)$  は第 k サブバンドの最低周波数、  $f_h(k)$  は第 k サブバンドの最高周波数を示し、 K はサブバンド数を示す。

そして、 Bark スケール形状算出器 507 は、各帯域の Bark スケール形状  $B(k)$  を量子化し、その符号化コードをマルチプレクサ 510 に送ると共に、 Bark スケール形状を復号化して Bark スケール正規化器 508 とベクトル量子化器 515 に与える。 Bark スケール正規化器 508 は、量子化後の Bark スケール形状を用いて正規化 M D C T 係数  $X3(m)$  を次の式 (13) に従い生成する。

$$X3(m) = \frac{X2(m)}{\sqrt{B_q(k)}} \quad f_l(k) \leq m \leq f_h(k) \quad 0 \leq k < K \quad \dots (13)$$

ここで  $Bq(k)$  は第  $k$  サブバンドの量子化後の Bark スケール形状を示す。

次に、ベクトル量子化器 509 では、Bark スケール正規化器 508 の出力  $X3(m)$  のベクトル量子化を行う。ベクトル量子化器 509 では、 $X3(m)$  を複数ベクトルに分割して各ベクトルに対応する符号帳を用いて最も歪が小さい代表値を求め、このインデックスを符号化コードとしてマルチプレクサ 510 に送る。

ベクトル量子化器 509 では、ベクトル量子化を行う際に 2 つの重要なパラメータを入力信号のスペクトル情報を用いて決定する。そのパラメータとは、1 つは量子化ビット配分であり、もう一つは符号帳探索時の重み付けである。量子化ビット配分は、スペクトル包絡算出器 502 で求められたスペクトル包絡  $env(m)$  を用いて決定する。

また、スペクトル包絡  $env(m)$  を用いて量子化ビット配分を決定する際に、周波数  $0 \sim F_L$  に相当するスペクトルに配分するビット数を小さくするように設定することもできる。

その一つの実現例として、周波数  $0 \sim F_L$  に配分できる最大ビット数  $MAX\_LOWBAND\_BIT$  を設定し、この帯域に配分されるビット数が最大ビット数  $MAX\_LOWBAND\_BIT$  を超えないように制限を設ける方法がある。

この実現例では、周波数  $0 \sim F_L$  では基本レイヤで符号化を既に行っているので、多くのビットを配分する必要がなく、この帯域での量子化を故意に粗くして、ビット配分を少なくし、そこで余分になるビットを周波数  $F_L \sim F_H$  に配分して量子化することにより全体的な品質を改善することができる。また、このビット配分は、スペクトル包絡  $env(m)$  と前述した Bark スケール形状  $Bq(k)$  とを組み合わせて決定する構成としても良い。

また、スペクトル包絡算出器 502 で求められたスペクトル包絡  $env(m)$  と Bark スケール形状算出器 507 で求められた量子化後の Bark スケール形状  $Bq(k)$  から算出される重み付けを利用した歪尺度を用いてベクトル量子化を行

う。ベクトル量子化は次の式（14）で規定される歪Dが最小となるコードベクトルCのインデックスjを求めて実現される。

$$D = \sum_m w(m)^2 (C_j(m) - X3(m))^2 \quad \dots (14)$$

5

ここで、w(m)は重み係数を示す。

また、重み関数w(m)はスペクトル包絡env(m)とBarkスケール形状Bq(k)を用いて次の式（15）のように表すことができる。

10  $w(m) = (env(m) \cdot Bq(Hertz\_to\_Bark(m)))^p \quad \dots (15)$

ここでpは0から1の間の定数、Herz\_to\_Bark()はHerzスケールをBarkスケールに変換する関数を示す。

また、重み関数w(m)を決定する際に、周波数0～FLに相当するスペクトルに配分する重み関数を小さくするように設定することも可能である。その一つの実現例として、周波数0～FLに対応する重み関数w(m)のとりうる最大値をMAX\_LOWBAND\_WGTとしてあらかじめ設定しておき、この帯域の重み関数w(m)の値がMAX\_LOWBAND\_WGTを超えないように制限を設ける方法がある。この実現例では、周波数0～FLでは基本レイヤで符号化を既に行っており、この帯域での量子化の精度を故意に下げて、相対的に周波数FL～FHの量子化の精度を上げることにより全体的な品質を改善することができる。

最後に、マルチプレクサ510では、符号化コードを多重化して、マルチプレクサ108に出力する。そして、新しい入力信号が存在する間、上記処理を

繰り返す。新しい入力信号が存在しない場合には、処理を終了する。

このように、本実施の形態の信号処理装置によれば、入力信号から所定の周波数以下の成分を取り出して符号励振線形予測法を用いた符号化を行い、得られた符号化コードを復号した結果を用いてMDCT変換で符号化を行うことによ

5 より、低ビットレートで高品質に符号化を行うことができる。

上記では、減算器106で得られる減算信号からLPC分析係数を分析している例について説明しているが、本発明の信号処理装置は、局所復号化器103において復号されたLPC係数を用いて符号化してもよい。

10 図6は、拡張レイヤ符号化器107の構成の一例を示す図である。但し、図5と同一の構成となるものについては、図5と同一番号を付し、詳しい説明を省略する。

15 図6の拡張レイヤ符号化器107は、変換テーブル601と、LPC係数マッピング部602と、スペクトル包絡算出器603と、変形部604とを具備し、局所復号化器103で復号されたLPC係数を用いて符号化する点が図5の拡張レイヤ符号化器107と異なる。

変換テーブル601は、基本レイヤのLPC係数と拡張レイヤのLPC係数とを対応づけて記憶する。

20 LPC係数マッピング部602は、変換テーブル601を参照し、局所復号化器103から入力された基本レイヤのLPC係数を拡張レイヤのLPC係数に変換し、スペクトル包絡算出器603に出力する。

スペクトル包絡算出器603は、拡張レイヤのLPC係数に基づいてスペクトル包絡を求め、変形部604に出力する。変形部604は、スペクトル包絡を変形し、スペクトル正規化器506及びベクトル量子化器509に出力する。

25 次に、図6の拡張レイヤ符号化器107の動作について説明する。基本レイヤのLPC係数は、信号帯域が0～FLの信号に対して求められたものであり、拡張レイヤの対象となる信号(信号帯域0～FH)で使用するLPC係数とは一

致しない。しかしながら、両者には強い相関がある。よってLPC係数マッピング部602では、この相関を利用してあらかじめ信号帯域0～FLの信号用のLPC係数と信号帯域0～FHの信号用のLPC係数との対応付けを表す変換テーブル601を別途設計しておく。この変換テーブル601を用いて、基

5 本レイヤのLPC係数から拡張レイヤのLPC係数を求める。

図7は、拡張LPC係数算出の一例を示す図である。変換テーブル601は、拡張レイヤのLPC係数(次数M)を表すJ個の候補 $\{Y_j(m)\}$ と、 $\{Y_j(m)\}$ と対応付けられた基本レイヤのLPC係数と同じ次数(=K)をもつ候補 $\{y_j(k)\}$ より構成される。 $\{Y_j(m)\}$ と $\{y_j(k)\}$ は大規模な楽音、音声データなどからあらかじめ設計して用意しておく。基本レイヤのLPC係数 $x(k)$ が入力されてきたとき、 $\{y_j(k)\}$ の中から $x(k)$ に最も類似しているLPC係数を求める。最も類似していると判定されたLPC係数のインデックスjに対応する拡張レイヤのLPC係数 $Y_j(m)$ を出力することにより、基本レイヤのLPC係数から拡張レイヤのLPC係数のマッピングを実現することができる。

15 次に、このようにして求めた拡張レイヤのLPC係数を基に、スペクトル包絡算出器603においてスペクトル包絡を求める。そして、このスペクトル包絡を変形部604において変形する。そして、この変形スペクトル包絡を前述した実施例のスペクトル包絡とみなして処理を行う。

スペクトル包絡を変形する変形部604の一つの実現例として、基本レイヤの符号化の対象となる信号帯域0～FLに対応するスペクトル包絡の影響を小さくする処理がある。スペクトル包絡を $env(m)$ としたとき、変形後のスペクトル包絡 $env'(m)$ は、以下の式(16)で表される。

$$env'(m) = \begin{cases} env(m)^p & \text{if } 0 \leq m \leq Fl \\ env(m) & \text{else} \end{cases} \dots (16)$$

ここで  $p$  は  $0 \sim 1$  の間の定数を示す。

- 。 周波数  $0 \sim F_L$  では基本レイヤで符号化を既に行っており、拡張レイヤの符号化対象である減算信号の周波数  $0 \sim F_L$  のスペクトルはフラットに近くなる。

5 それにも関わらず、本実施例で説明したような L P C 係数のマッピングではこのような作用は考慮されていない。そこで、式 (16) を用いてスペクトル包絡を修正する手法を用いることにより品質改善を図ることができる。

10 このように、本実施の形態の信号処理装置によれば、基本レイヤ符号化器で量子化した L P C 係数を用いて拡張レイヤの L P C 係数を求め、拡張レイヤの L P C 分析からスペクトル包絡を算出することより、L P C 分析および量子化の必要がなくなり、量子化ビット数を削減することができる。

(実施の形態 3)

15 図 8 は、本発明の実施の形態 3 に係る信号処理装置の拡張レイヤ符号化器の構成を示すブロック図である。但し、図 5 と同一の構成となるものについては、図 5 と同一番号を付し、詳しい説明を省略する。

20 図 8 の拡張レイヤ符号化器 107 は、スペクトル微細構造算出器 801 を具備し、基本レイヤ符号化器 102 で符号化され局所復号化器 103 で復号されたピッチ周期を用いてスペクトル微細構造を算出し、当該スペクトル微細構造をスペクトルの正規化およびベクトル量子化に活用する点が図 5 の拡張レイヤ符号化器と異なる。

スペクトル微細構造算出器 801 は、基本レイヤで符号化されたピッチ周期  $T$  とピッチゲイン  $\beta$  からスペクトル微細構造を算出し、スペクトル正規化器 5

0 6 に出力する。

具体的には、これらピッチ周期  $T$  とピッチゲイン  $\beta$  は符号化コードの一部で  
あり、ここでは図示されない音響復号器において同じ情報を得ることができる。  
よって、ピッチ周期  $T$  とピッチゲイン  $\beta$  を利用して符号化を行ったとしてもビ  
5 ットレートが増加することはない。

スペクトル微細構造算出器 801 では、ピッチ周期  $T$  とピッチゲイン  $\beta$  を用  
いて次の式 (17) に従いスペクトル微細構造  $har(m)$  を算出する。

$$har(m) = \left| \frac{1}{1 - \beta \cdot e^{-j \frac{2\pi m T}{M}}} \right| \quad \dots (17)$$

10

ここで  $M$  はスペクトル分解能を示す。式 (17) は、 $\beta$  の絶対値が 1 以上となる場合に発振フィルタとなるため、 $\beta$  の絶対値がとりうる範囲をあらかじめ定められた 1 未満の設定値 (例えば 0.8) 以下になるよう制限を設ける方法もある。

15 スペクトル正規化器 506 では、スペクトル包絡算出器 502 で求められるスペクトル包絡  $env(m)$  と、スペクトル微細構造算出器 801 で求められるスペクトル微細構造  $har(m)$  の両者を用いて次の式 (18) に従い正規化を行う。

$$X2(m) = \frac{X1(m)}{env(m) \cdot har(m)} \quad \dots (18)$$

また、ベクトル量子化器 509 での量子化ビットの配分は、スペクトル包絡算出器 502 で求められるスペクトル包絡  $env(m)$  とスペクトル微細構造算出器 5 801 で求められるスペクトル微細構造  $har(m)$  の両者を用いて決定する。また、ベクトル量子化の際の重み関数  $w(m)$  の決定に、スペクトル微細構造をも利用する。具体的には、重み関数  $w(m)$  は次の式 (19) に従い定義される。

$$w(m) = (env(m) \cdot har(m) \cdot Bq(Hertz\_to\_Bark(m)))^p \quad \dots (19)$$

10

ここで  $p$  は 0 から 1 の間の定数、 $Herz\_to\_Bark()$  は Herz スケールを Bark スケールに変換する関数を示す。

このように、本実施の形態の信号処理装置は、基本レイヤ符号化器で符号化され局所復号化器で復号されたピッチ周期を用いてスペクトル微細構造を算出し、当該スペクトル微細構造をスペクトルの正規化およびベクトル量子化に活用することにより、量子化性能を向上することができる。

#### (実施の形態 4)

図 9 は、本発明の実施の形態 4 に係る信号処理装置の拡張レイヤ符号化器の構成を示すブロック図である。但し、図 5 と同一の構成となるものについては、20 図 5 と同一番号を付し、詳しい説明を省略する。

図 9 の拡張レイヤ符号化器 107 は、パワー推定器 901 と、パワー変動量量子化器 902 とを具備し、基本レイヤ符号化器 102 により得られる符号化

コードを用いて局所復号化器 103において復号信号を生成し、当該復号信号からMDCT係数のパワーを予測し、その予測値からの変化量を符号化する点が図5の拡張レイヤ符号化器と異なる。

また図1において局所復号化器103から拡張レイヤ107に復号されたパラメータが出力されているが、本実施の形態では復号パラメータの代わりに局所復号化器103において求められる復号信号が拡張レイヤ符号化器107に出力される。

図5における局所復号化器103において復号化された信号  $sl(n)$  が、パワー推定器901に入力される。そして、パワー推定器901では、前記復号信号  $sl(n)$  からMDCT係数のパワーを推定する。MDCT係数のパワーの推定値を  $powp$  とすると、  $powp$  は次の式(20)のように表される。

$$powp = \alpha \cdot \sum_{n=0}^{N-1} sl(n)^2 \quad \dots (20)$$

ここで  $N$  は復号信号  $sl(n)$  の長さ、  $\alpha$  はあらかじめ定められた補正のための定数を示す。また、基本レイヤのLPC係数から求められるスペクトル傾きを利用した別の方法では、MDCT係数のパワーの推定値は次の式(21)にて表される。

$$powp = \alpha \cdot \beta \cdot \sum_{n=0}^{N-1} sl(n)^2 \quad \dots (21)$$

ここで  $\beta$  は、基本レイヤのLPC係数から求められるスペクトル傾きに依存し

た変数を表し、スペクトル傾きが大きい(相対的に低域にパワーがある)場合に  
 $\beta$  はゼロに近づき、スペクトル傾きが小さい(相対的に高域にパワーがある)場合に  $\beta$  は 1 に近づく性質を持つ。

次に、パワー変動量量子化器 902 では、MCDT 部 503 で求めた MCDT  
5 係数のパワーをパワー推定器 901 で求めたパワー推定値  $pow_p$  にて正規化し、  
その変動量を量子化する。変動量  $r$  は次の式 (22) で表される。

$$r = \frac{pow}{pow_p} \quad \dots (22)$$

10 ここで  $pow$  は MCDT 係数のパワーを示し、式 (23) にて算出される。

$$pow = \sum_{m=0}^{M-1} X(m)^2 \quad \dots (23)$$

ここで  $X(m)$  は MCDT 係数、 $M$  はフレーム長を示す。パワー変動量量子化器 9  
15 02 では、変動量  $r$  を量子化し、その符号化コードをマルチプレクサ 510 に  
送ると共に、量子化後の変動量  $r_q$  を復号する。パワー正規化器 505 では、  
量子化後の変動量  $r_q$  を用いて MCDT 係数を次の式 (24) を用いて正規化  
する。

$$X1(m) = \frac{X(m)}{\sqrt{rq \cdot powp}} \quad \dots (24)$$

ここで、 $X1(m)$ はパワー正規化後のMDCT係数を示す。

このように、本実施の形態の信号処理装置は、基本レイヤの復号信号のパワーと拡張レイヤのMDCT係数のパワーとの間の相関を利用し、基本レイヤの復号信号を利用してMDCT係数のパワーを予測し、その予測値からの変動量を符号化することにより、MDCT係数のパワーの量子化に必要なビット数を削減することができる。

(実施の形態 5)

図10は、本発明の実施の形態5に係る信号処理装置の構成を示すブロック図である。図10の信号処理装置1000は、デマルチプレクサ1001と、基本レイヤ復号化器1002と、アップサンプリング器1003と、拡張レイヤ復号化器1004と、加算器1005とから主に構成される。

デマルチプレクサ1001は、符号化されたコードを分離して基本レイヤ用の符号化コードと拡張レイヤ用の符号化コードを生成する。そして、デマルチプレクサ1001は、基本レイヤ用の符号化コードを基本レイヤ復号化器1002に出力し、拡張レイヤ用の符号化コードを拡張レイヤ復号化器1004に出力する。

基本レイヤ復号化器1002は、デマルチプレクサ1001で得られた基本レイヤ用の符号化コードを用いてサンプリングレートFLの復号信号を復号し、アップサンプリング器1003に出力する。同時に、基本レイヤ復号化器1002で復号されたパラメータを拡張レイヤ復号化器1004に出力する。アップサンプリング器1003は、復号信号のサンプリング周波数をFHに上げ、加算器1005に出力する。

拡張レイヤ復号化器 1004 は、デマルチプレクサ 1001 で得られた拡張  
レイヤ用の符号化コードと基本レイヤ復号化器 1002 において復号されたパ  
ラメータを用いてサンプリングレート FH の復号信号を復号し、加算器 100  
5 に出力する。

5 加算器 1005 は、アップサンプリング器 1003 から出力された復号信号  
と、拡張レイヤ復号化器 1004 から出力された復号信号をベクトル加算する。

次に、本実施の形態の信号処理装置の動作について説明する。最初に、実施  
の形態 1 から 4 のいずれかの信号処理装置において符号化されたコードが入力  
され、デマルチプレクサ 1001 において前記コードを分離して基本レイヤ用  
10 の符号化コードと拡張レイヤ用の符号化コードを生成する。

次に、基本レイヤ復号化器 1002 では、デマルチプレクサ 1001 で得ら  
れた基本レイヤ用の符号化コードを用いてサンプリングレート FL の復号信号  
を復号する。そして、アップサンプリング器 1003 は、当該復号信号をサン  
プリング周波数を FH に上げる。

15 拡張レイヤ復号化器 1004 では、デマルチプレクサ 1001 で得られた拡  
張レイヤ用の符号化コードと基本レイヤ復号化器 1002 において復号された  
パラメータを用いてサンプリングレート FH の復号信号が復号される。

前記アップサンプリング器 1003 においてアップサンプリングされた基本  
レイヤの復号信号と当該拡張レイヤの復号信号とを加算器 1005 において加  
20 算する。そして、新しい入力信号が存在する間、上記処理を繰り返す。新しい  
入力信号が存在しない場合には、処理を終了する。

このように、本実施の形態の信号処理装置は、基本レイヤ復号化器 1002  
で復号されたパラメータを用いて拡張レイヤ復号化器 1004 の復号を行うこ  
とにより、基本レイヤ符号化における復号パラメータを使って拡張レイヤの符  
25 号化を行う音響符号化手段の符号化コードから復号信号を生成することができ  
る。

次に、基本レイヤ復号化器 1002について説明する。図11は、基本レイヤ復号化器 1002の一例を示すブロック図である。図11の基本レイヤ復号化器 1002は、デマルチプレクサ 1101と、音源生成器 1102と、合成フィルタ 1103とから主に構成され、CELPの復号化処理を行う。

5 デマルチプレクサ 1101は、デマルチプレクサ 1001から出力された基本レイヤ用の符号化コードから各種パラメータを分離し、音源生成器 1102及び合成フィルタ 1103に出力する。

音源生成器 1102は、適応ベクトル、適応ベクトルゲイン、雑音ベクトル、  
10 雜音ベクトルゲインを復号し、これらを用いて音源信号を生成し合成フィルタ 1103に出力する。合成フィルタ 1103は、復号されたLPC係数を用いて合成信号を生成する。

次に、図11の基本レイヤ復号化器 1002の動作について説明する。最初に、デマルチプレクサ 1101は、基本レイヤ用の符号化コードから、各種パラメータを分離する。

15 次に、音源生成器 1102が、適応ベクトル、適応ベクトルゲイン、雑音ベクトル、雑音ベクトルゲインを復号する。そして、音源生成器 1102は、次の式 (25) に従い音源ベクトル  $ex(n)$  を生成する。

$$ex(n) = \beta_q \cdot q(n) + \gamma_q \cdot c(n) \quad \dots (25)$$

20 ここで、 $q(n)$  は適応ベクトル、 $\beta_q$  は適応ベクトルゲイン、 $c(n)$  は雑音ベクトル、 $\gamma_q$  は雑音ベクトルゲインを示す。

次に、合成フィルタ 1103が、復号されたLPC係数を用いて合成信号  $syn(n)$  を次の式 (26) に従い生成する。

$$syn(n) = ex(n) + \sum_{i=1}^{NP} \alpha_q(i) \cdot syn(n-i) \quad \dots (26)$$

ここで、 $\alpha_q$ は復号されたLPC係数、NPはLPC係数の次数を示す。

このように復号された復号信号  $syn(n)$  は、アップサンプリング器 1003 に  
 5 出力し、復号された結果求められるパラメータを拡張レイヤ復号化器 1004 に  
 に出力される。そして、新しい入力信号が存在する間、上記処理を繰り返す。  
 新しい入力信号が存在しない場合には、処理を終了する。CELP の構成によつ  
 ては、合成信号をポストフィルタに通した後に出力する形態もありうる。ここ  
 でいうポストフィルタとは、符号化歪を知覚しにくくする後処理の機能を有す  
 10 るものである。

次に、拡張レイヤ復号化器 1004 について説明する。図 12 は、拡張レイ  
 ヤ復号化器 1004 の一例を示すブロック図である。図 12 の拡張レイヤ復号  
 化器 1004 は、デマルチプレクサ 1201 と、LPC 係数復号化器 1202  
 と、スペクトル包絡算出器 1203 と、ベクトル復号化器 1204 と、Bark ス  
 15 ケール形状復号化器 1205 と、乗算器 1206 と、乗算器 1207 と、パワ  
 一復号化器 1208 と、乗算器 1209 と、IMDCT 部 1210 とから主に  
 構成される。

デマルチプレクサ 1201 は、デマルチプレクサ 1001 から出力された拡  
 張レイヤ用の符号化コードから各種パラメータを分離する。LPC 係数復号化  
 器 1202 は、LPC 係数に関する符号化コードを用いて LPC 係数を復号し、  
 20 スペクトル包絡算出器 1203 に出力する。

スペクトル包絡算出器 1203 は、復号された LPC 係数を用いて式 (6)  
 に従いスペクトル包絡  $env(m)$  を算出し、ベクトル復号化器 1204 及び乗算器  
 1207 に出力する。

ベクトル復号化器 1204 は、スペクトル包絡算出器 1203 において求められたスペクトル包絡  $env(m)$  に基づいて量子化ビット配分を決定し、デマルチプレクサ 1201 から得られる符号化コードと前記量子化ビット配分とから正規化MDCT係数  $X3q(m)$  を復号する。なお、量子化ビット配分の方法は、実施 5 の形態 1 から実施の形態 4 のいずれかの符号化法において拡張レイヤ符号化で用いたものと同じ方法とする。

Bark スケール形状復号化器 1205 は、デマルチプレクサ 1201 より得られる符号化コードを元に Bark スケール形状  $Bq(k)$  を復号し、乗算器 1206 に 10 出力する。

乗算器 1206 は、次の式 (27) に従い、正規化MDCT係数  $X3q(m)$  と Bark スケール形状  $Bq(k)$  を乗算し、乗算結果を乗算器 1207 に出力する。

$$X2_q(m) = X3_q(m) \sqrt{B_q(k)} \quad f_l(k) \leq m \leq f_h(k) \quad 0 \leq k < K \quad \cdots (27)$$

15

ここで  $f_l(k)$  は第  $k$  サブバンドの最低周波数、 $f_h(k)$  は第  $k$  サブバンドの最高周波数を表し、 $K$  はサブバンド数を示す。

乗算器 1207 は、次の式 (28) に従い、乗算器 1206 より得られる正規化MDCT係数  $X2q(m)$  とスペクトル包絡算出器 1203 において求められたスペクトル包絡  $env(m)$  を乗算し、乗算結果を乗算器 1209 に出力する。

$$X1_q(m) = X2_q(m) env(m) \quad \cdots (28)$$

25 パワー復号化器 1208 は、デマルチプレクサ 1201 より得られる符号化コードを元にパワー  $powq$  を復号し、復号結果を乗算器 1209 に出力する。

乗算器 1209 は、次の式 (29) に従い、正規化MDCT係数  $X1q(m)$  と復号パワー  $powq$  を乗算し、乗算結果を IMDCT 部 1210 に出力する。

5       $X_q(m) = X1_q(m) \sqrt{powq} \quad \cdots (29)$

IMDCT 部 1210 は、このようにして求められた復号MDCT係数に IMDCT 変換 (Inverse Modified Discrete Cosine Transform: 逆修正コサイン変換) を施し、前フレームで復号された信号と分析フレームの半分だけオーバーラップさせて加算して出力信号を生成し、この出力信号を加算器 1005 に出力する。そして、新しい入力信号が存在する間、上記処理を繰り返す。新しい入力信号が存在しない場合には、処理を終了する。

このように、本実施の形態の信号処理装置によれば、基本レイヤ復号化器で復号されたパラメータを用いて拡張レイヤ復号化器の復号を行うことにより、15 基本レイヤ符号化における復号パラメータを使って拡張レイヤの符号化を行う音響符号化手段の符号化コードから復号信号を生成することができる。

(実施の形態 6)

図 13 は、拡張レイヤ復号化器 1004 の構成の一例を示す図である。但し、図 12 と同一の構成となるものについては、図 12 と同一番号を付し、詳しい20 説明を省略する。

図 13 の拡張レイヤ復号化器 1004 は、変換テーブル 1301 と、LPC 係数マッピング部 1302 と、スペクトル包絡算出器 1303 と、変形部 1304 とを具備し、基本レイヤ復号化器 1002 で復号された LPC 係数を用いて復号化する点が図 12 の拡張レイヤ復号化器 1004 と異なる。

25      変換テーブル 1301 は、基本レイヤの LPC 係数と拡張レイヤの LPC 係数とを対応づけて記憶する。

LPC係数マッピング部1302は、変換テーブル1301を参照し、基本レイヤ復号化器1002から入力された基本レイヤのLPC係数を拡張レイヤのLPC係数に変換し、スペクトル包絡算出器1303に出力する。

スペクトル包絡算出器1303は、拡張レイヤのLPC係数に基づいてスペクトル包絡を求め、変形部1304に出力する。変形部1304は、スペクトル包絡を変形し、乗算器1207及びベクトル復号器1204に出力する。例えば、変形の方法は、実施の形態2の式(16)で示される方法がある。

次に、図13の拡張レイヤ復号化器1004の動作について説明する。基本レイヤのLPC係数は、信号帯域が0～FLの信号に対して求められたものであり、拡張レイヤの対象となる信号(信号帯域0～FH)で使用するLPC係数とは一致しない。しかしながら、両者には強い相関がある。よってLPC係数マッピング部1302では、この相関を利用してあらかじめ信号帯域0～FLの信号用のLPC係数と信号帯域0～FHの信号用のLPC係数との対応付けを表す変換テーブル1301を別途設計しておく。この変換テーブル1301を用いて、基本レイヤのLPC係数から拡張レイヤのLPC係数を求める。

変換テーブル1301の詳細は、実施の形態2の変換テーブル601と同様である。

このように、本実施の形態の信号処理装置によれば、基本レイヤ復号化器で量子化したLPC係数を用いて拡張レイヤのLPC係数を求め、拡張レイヤのLPC係数からスペクトル包絡を算出することより、LPC分析および量子化の必要がなくなり、量子化ビット数を削減することができる。

#### (実施の形態7)

図14は、本発明の実施の形態7に係る信号処理装置の拡張レイヤ復号化器の構成を示すブロック図である。但し、図12と同一の構成となるものについては、図12と同一番号を付し、詳しい説明を省略する。

図14の拡張レイヤ復号化器1004は、スペクトル微細構造算出器140

1を具備し、基本レイヤ復号化器1002で復号化されたピッチ周期を用いてスペクトル微細構造を算出し、当該スペクトル微細構造を復号化に活用し、量子化性能を向上させた音響符号化に対応した音響復号化を行う点が図12の拡張レイヤ符号化器と異なる。

5 スペクトル微細構造算出器1401は、基本レイヤ復号化器1002で復号化されたピッチ周期Tとピッチゲイン $\beta$ からスペクトル微細構造を算出し、ベクトル復号化器1204及び乗算器1207に出力する。

スペクトル微細構造算出器1401では、ピッチ周期 $T_q$ とピッチゲイン $\beta_q$ を用いて次の式(17)に従いスペクトル微細構造 $har(m)$ を算出する。

10

$$har(m) = \left| \frac{1}{1 - \beta_q \cdot e^{-j \frac{2\pi m T_q}{M}}} \right| \quad \dots (30)$$

ここでMはスペクトル分解能を示す。式(17)は、 $\beta_q$ の絶対値が1以上となる場合に発振フィルタとなるため、 $\beta_q$ の絶対値がとりうる範囲をあらかじめ定められた1未満の設定値(例えば0.8)以下になるよう制限を設けてよい。

そして、スペクトル包絡算出器1203で求められたスペクトル包絡 $env(m)$ とスペクトル微細構造算出器1401で求められたスペクトル微細構造 $har(m)$ を用いて、ベクトル復号化器1204での量子化ビット配分が決定される。そ

して、当該量子化ビット配分とデマルチプレクサ 1201 から得られる符号化コードから正規化MDCT係数  $X3q(m)$  が復号される。さらに、乗算器 1207 において次の式 (30) に従い、正規化MDCT係数  $X2q(m)$  にスペクトル包絡  $env(m)$  とスペクトル微細構造  $har(m)$  を乗じて正規化MDCT係数  $X1q(m)$  が求められる。

$$X1_q(m) = X2_q(m)env(m)har(m) \quad \dots (31)$$

このように、本実施の形態の信号処理装置は、基本レイヤ符号化器で符号化され局所復号化器で復号されたピッチ周期を用いてスペクトル微細構造を算出し、当該スペクトル微細構造をスペクトルの正規化およびベクトル量子化に活用することにより、量子化性能を向上させた音響符号化に対応した音響復号化を行うことができる。

(実施の形態 8)

図 15 は、本発明の実施の形態 8 に係る信号処理装置の拡張レイヤ復号化器の構成を示すブロック図である。但し、図 12 と同一の構成となるものについては、図 12 と同一番号を付し、詳しい説明を省略する。

図 15 の拡張レイヤ復号化器 1004 は、パワー推定器 1501 と、パワー変化量復号化器 1502 と、パワー生成器 1503 とを具備し、基本レイヤの復号信号を利用してMDCT係数のパワーを予測し、その予測値からの変化量を符号化する符号化器に対応する復号化器を構成している点が図 12 の拡張レイヤ復号化器と異なる。

また、図 10 において、基本レイヤ復号化器 1002 から拡張レイヤ復号化器 1004 に復号されたパラメータが出力されているが、本実施の形態では、

復号パラメータの代わりに基本レイヤ復号化器 1002において得られる復号信号が拡張レイヤ復号化器 1004に出力される。

パワー推定器 1501 は、基本レイヤ復号化器 1002において復号化された復号信号  $s_l(n)$  から MDCT 係数のパワーを式 (20) または式 (21) を用いて推定する。

パワー変化量復号化器 1502 では、デマルチプレクサ 1201 から得られる符号化コードからパワー変化量を復号し、パワー生成器 1503 に出力する。

パワー生成器 1503 は、パワー変化量からパワーを算出する。

乗算器 1209 は、次の式 (31) に従い MDCT 係数を求める。

10

$$X_q(m) = X1_q(m) \sqrt{rq \cdot powp} \quad \dots (32)$$

ここで、 $rq$  はパワー変化量の復号値、 $powp$  はパワー推定値を示す。また、 $X1q(m)$  は乗算器 1207 の出力信号を示す。

このように、本実施の形態の信号処理装置によれば、基本レイヤの復号信号を利用して MDCT 係数のパワーを予測し、その予測値からの変化量を符号化する符号化器に対応する復号化器を構成していることにより、MDCT 係数のパワーの量子化に必要なビット数を削減することができる。

20 (実施の形態 9)

図 16 は、本発明の実施の形態 9 に係る音響符号化装置の構成を示すブロック図である。図 16 の音響符号化装置 1600 は、ダウンサンプリング器 1601 と、基本レイヤ符号化器 1602 と、局所復号化器 1603 と、アップサンプリング器 1604 と、遅延器 1605 と、減算器 1606 と、周波数決定部 1607 と、拡張レイヤ符号化器 1608 と、多重化器 1609 とから主に構成される。

図16において、ダウンサンプリング器1601は、サンプリングレートFHの入力データ（音響データ）を受けつけ、この入力データをサンプリングレートFHより低いサンプリングレートFLに変換して基本レイヤ符号化器1602に出力する。

5 基本レイヤ符号化器1602は、サンプリングレートFLの入力データを所定の基本フレーム単位で符号化し、入力データを符号化した第1符号化コードを局所復号化器1603と多重化器1609に出力する。例えば、基本レイヤ符号化器1602は、入力データをCELP方式で符号化する。

10 局所復号化器1603は、第1符号化コードを復号化し、復号化により得られた復号信号をアップサンプリング器1604に出力する。アップサンプリング器1604は、復号信号のサンプリングレートをFHに上げて減算器1606と周波数決定部1607に出力する。

15 遅延器1605は、入力信号を所定の時間遅延して減算器1606に出力する。この遅延の大きさをダウンサンプリング器1601と基本レイヤ符号化器1602と局所復号化器1603とアップサンプリング器1604で生じる時間遅れと同値とすることにより、次の減算処理での位相のずれを防ぐ役割を持つ。減算器1606は、入力信号を復号信号で減算し、減算結果を誤差信号として拡張レイヤ符号化器1608に出力する。

20 周波数決定部1607は、サンプリングレートをFHに上げた復号信号から誤差信号の符号化する領域と、符号化しない領域を決定して拡張レイヤ符号化器1608に通知する。例えば、周波数決定部1607は、サンプリングレートをFHに上げた復号信号から聴覚マスキングする周波数を決定して拡張レイヤ符号化器1608に出力する。

25 拡張レイヤ符号化器1608は、誤差信号を周波数領域の係数に変換して誤差スペクトルを生成し、周波数決定部1607から得られる符号化の対象となる周波数情報に基づき誤差スペクトルの符号化を行う。多重化器1609は、

基本レイヤ符号化器1602で符号化して求められる符号化コードと、拡張レイヤ符号化器1608で符号化して求められる符号化コードを多重化する。

以下、基本レイヤ符号化器1602と拡張レイヤ符号化器1608とがそれぞれ符号化する信号について説明する。図17は、音響信号の情報の分布の一例を示す図である。図17において、縦軸は情報量を示し、横軸は周波数を示す。図17では、入力信号に含まれる音声情報と背景音楽・背景雑音情報がどの周波数帯にどれだけ存在しているかを表している。

図17に示すように、音声情報は、周波数の低い領域に情報が多く存在し、高域に向かうほど情報量は減少する。一方、背景音楽・背景雑音情報は、音声情報と比べると相対的に低域の情報は少なく、高域に含まれる情報が大きい。

そこで、基本レイヤではC E L Pを用いて音声信号を高品質に符号化し、拡張レイヤでは基本レイヤで表しきれない背景の音楽や環境音、基本レイヤでカバーする周波数帯よりも高い周波数成分の信号を効率よく符号化する。

図18は、基本レイヤと拡張レイヤで符号化の対象とする領域の一例を示す図である。図18において、縦軸は情報量を示し、横軸は周波数を示す。図18は、基本レイヤ符号化器1602と拡張レイヤ符号化器1608がそれぞれ符号化する情報の対象となる領域を表している。

基本レイヤ符号化器1602は、0～F L間の周波数帯の音声情報を効率よく表すように設計されており、この領域での音声情報は品質良く符号化することができる。しかし、基本レイヤ符号化器1602では、0～F L間の周波数帯の背景音楽・背景雑音情報の符号化品質が高くない。

拡張レイヤ符号化器1608は、上記説明にある基本レイヤ符号化器1602の能力不足の部分と、F L～F H間の周波数帯の信号をカバーするように設計されている。よって、基本レイヤ符号化器1602と拡張レイヤ符号化器1608を組み合わせることで広い帯域で高品質な符号化が実現できる。

図18に示すように、基本レイヤ符号化器1602における符号化により得

られた第1符号化コードには、0～FL間の周波数帯の音声情報が含まれているので、少なくとも第1符号化コードのみでも復号信号が得られるというスケーラブル機能が実現できる。

また、拡張レイヤで聴覚マスキングを利用して符号化効率を上げることが考  
5 えられる。聴覚マスキングとは、ある信号が与えられたときその信号の周波数  
の近傍に位置する信号が聞こえなくなる（マスクされる）という人間の聴覚特  
性を利用したものである。

図19は、音響（音楽）信号のスペクトルの一例を示す図である。図19に  
おいて、実線は聴覚マスキングを表し、破線は誤差スペクトルを表している。  
10 ここでいう誤差スペクトルとは、入力信号と基本レイヤの復号信号との誤差信  
号（拡張レイヤの入力信号）のスペクトルを指す。

図19の斜線部で表される誤差スペクトルは、聴覚マスキングよりも振幅値  
が小さいため人間の聴覚では聞こえず、それ以外の領域では誤差スペクトルの  
振幅値が聴覚マスキングを超えてるので量子化歪が知覚される。

15 そこで、拡張レイヤでは図19の白地部に含まれる誤差スペクトルを符号化  
してその領域の量子化歪が聴覚マスキングよりも小さくなるようにすればよい。  
また、斜線部に属する係数は既に聴覚マスキングよりも小さいので量子化する  
必要がない。

本実施の形態の音響符号化装置1600では、聴覚マスキング等により残差  
20 信号を符号化する周波数を符号化側から復号化側に伝送することをせず、符号  
化側と復号側でそれぞれアップサンプリングされた基本レイヤの復号信号を用  
いて拡張レイヤが符号化する誤差スペクトルの周波数を決定する。

25 基本レイヤの符号化コードを復号化した復号信号は、符号化側と復号化側で  
同じ信号が得られるので、符号化側は、この復号化信号から聴覚マスキングす  
る周波数を決定して信号を符号化し、復号化側は、この復号化信号から聴覚マ  
スキングされた周波数の情報を得て信号を復号化することにより、誤差スペク

トルの周波数の情報を付加情報として符号化して伝送する必要は無くなり、ビットレートの低減を実現することができる。

次に、本実施の形態に係る音響符号化装置の各ブロックの詳細な動作について説明する。最初にアップサンプリングされた基本レイヤの復号信号（以後、  
5 基本レイヤ復号信号と呼ぶ）から拡張レイヤにて符号化される誤差スペクトルの周波数を決定する周波数決定部 1607 の動作の説明を行う。図 20 は、本実施の形態の音響符号化装置の周波数決定部の内部構成の一例を示すブロック図である。

図 20において、周波数決定部 1607 は、FFT 部 1901 と、推定聴覚マスキング算出器 1902 と、決定部 1903 とから主に構成される。  
10

FFT 部 1901 は、アップサンプリング器 1604 から出力された基本レイヤ復号信号  $x(n)$  を直交変換して振幅スペクトル  $P(m)$  を算出して推定聴覚マスキング算出器 1902 と決定部 1903 に出力する。具体的には、FFT 部 1901 は、以下の式（33）を用いて振幅スペクトル  $P(m)$  を算出  
15 する。

$$P(m) = \sqrt{\operatorname{Re}^2(m) + \operatorname{Im}^2(m)} \quad \dots (33)$$

ここで、 $\operatorname{Re}(m)$  と  $\operatorname{Im}(m)$  は基本レイヤ復号信号  $x(n)$  のフーリエ係数の実部と虚部、 $m$  は周波数を表す。  
20

次に、推定聴覚マスキング算出器 1902 は、基本レイヤ復号信号の振幅スペクトル  $P(m)$  を用いて推定聴覚マスキング  $M'(m)$  を算出して決定部 1903 に出力する。一般的には、聴覚マスキングは、入力信号のスペクトルを

基に算出されるものであるが、本実施例では入力信号の代わりに基本レイヤ復号信号  $x(n)$  を使って聴覚マスキングを推定する。これは、基本レイヤ復号信号  $x(n)$  は入力信号との歪が小さくなるよう決定されているため、入力信号の代わりに基本レイヤ復号信号  $x(n)$  を用いても充分に近似され大きな問題は生じないという考えに基づいている。

5 次に、決定部 1903 は、基本レイヤ復号信号の振幅スペクトル  $P(m)$  と推定聴覚マスキング算出器 1902 で得られる推定聴覚マスキング  $M'(m)$  を用いて拡張レイヤ符号化器 1608 で誤差スペクトルを符号化する対象の周波数を決定する。決定部 1903 は、基本レイヤ復号信号の振幅スペクトル  $P(m)$  を誤差スペクトルの近似値とみなし、次の式 (34) が成立する。

10  $m$  を拡張レイヤ符号化器 1608 に出力する。

$$P(m) - M'(m) > 0 \quad \dots (34)$$

15 式 (34)において、 $P(m)$  の項は、誤差スペクトルの大きさを推定しており、 $M'(m)$  の項は、聴覚マスキングを推定している。そして、決定部 1903 は、推定誤差スペクトルと推定聴覚マスキングの大きさを比較し、式 (34) を満たす場合、すなわち推定聴覚マスキングの大きさを推定誤差スペクトルの大きさが超える場合に、その周波数の誤差スペクトルはノイズとして知覚されるとして拡張レイヤ符号化器 1608 で符号化する対象とする。

20 逆に推定聴覚マスキングの大きさより推定誤差スペクトルの大きさが下回る場合に、決定部 1903 は、マスキング効果によりその周波数の誤差スペクトルはノイズとして知覚されないとみなし、この周波数の誤差スペクトルは量子

化の対象から外す。

次に、推定聴覚マスキング算出器 1902 の動作を説明する。図 21 は、本実施の形態の音響符号化装置の聴覚マスキング算出器の内部構成の一例を示す図である。図 21 において、推定聴覚マスキング算出器 1902 は、バーカスペクトル算出器 2001 と、スプレッド関数畳み込み器 2002 と、トーナリティ算出器 2003 と、聴覚マスキング算出器 2004 とから主に構成される。

図 21 において、バーカスペクトル算出器 2001 は、以下の式 (35) を用いてバーカスペクトル  $B(k)$  を算出する。

$$B(k) = \sum_{m=fL(k)}^{fH(k)} P^2(m) \quad \dots (35)$$

ここで、 $P(m)$  は振幅スペクトルを表し、上述の式 (33) より求められる。また、 $k$  はバーカスペクトルの番号に対応し、 $F_L(k)$ 、 $F_H(k)$  はそれぞれ第  $k$  バーカスペクトルの最低周波数、最高周波数を表す。バーカスペクトル  $B(k)$  はバーカスケール上で等間隔に帯域分割されたときのスペクトル強度を表す。ヘルツスケールを  $f$ 、バーカスケールを  $B$  と表したとき、ヘルツスケールとバーカスケールの関係は以下の式 (36) で表される。

$$B = 13 \tan^{-1}(0.76f) + 3.5 \tan^{-1}\left(\frac{f}{7.5}\right) \quad \dots (36)$$

スプレッド関数畳み込み器2002は、以下に示す式(37)を用いてバーカスペクトルB(k)にスプレッド関数SF(k)を畳み込み、C(k)を算出する。

$$C(k) = B(k) * SF(k) \quad \dots (37)$$

トーナリティ算出器2003は、以下の式(38)を用い、各バーカスペクトルのスペクトル平坦度SFM(k)を求める。

$$SFM(k) = \frac{\mu g(k)}{\mu a(k)} \quad \dots (38)$$

ここで、 $\mu_g(k)$  は第  $k$  バークスペクトルに含まれるパワースペクトルの幾何平均、 $\mu_a(k)$  は第  $k$  バークスペクトルに含まれるパワースペクトルの算術平均を表す。そして、トーナリティ算出器 2003 は、以下の式 (39) を用いてスペクトル平坦度  $S_{FM}(k)$  のデシベル値  $S_{FMdB}(k)$  からトーナリティ係数  $\alpha(k)$  を算出する。

$$\alpha(k) = \min\left(\frac{S_{FMdB}(k)}{-60}, 1.0\right) \quad \dots (39)$$

聴覚マスキング算出器 2004 は、以下の式 (40) を用いてトーナリティ算出器 2003 で算出したトーナリティ係数  $\alpha(k)$  から各バークスケールのオフセット  $O(k)$  を求める。

$$O(k) = \alpha(k) \cdot (14.5 - k) + (1.0 - \alpha(k)) \cdot 5.5 \quad \dots (40)$$

そして、聴覚マスキング算出器 2004 は、以下の式 (41) を用いてスプレッド関数畳み込み器 2002 で求めた  $C(k)$  からオフセット  $O(k)$  を減算して聴覚マスキング  $T(k)$  を算出する。

$$T(k) = \max\left(10^{\log_{10}(C(k)) - (O(k)/10)}, T_q(k)\right) \dots (41)$$

ここで、 $T_q(k)$  は絶対閾値を表す。絶対閾値は、人間の聴覚特性として観測される聴覚マスキングの最小値を表す。そして、聴覚マスキング算出器 20 5 0 4 は、バーカスケールで表される聴覚マスキング  $T(k)$  をヘルツスケールに変換して推定聴覚マスキング  $M'(m)$  を求め、決定部 1903 に出力する。

このようにして求められた量子化の対象となる周波数  $m$  を使って、拡張レイヤ符号化器 1608 にて MDCT 係数の符号化を行う。図 22 は、本実施の形態の拡張レイヤ符号化器の内部構成の一例を示すブロック図である。図 22 の 10 拡張レイヤ符号化器 1608 は、MDCT 部 2101 と、MDCT 係数量子化器 2102 とから主に構成される。

MDCT 部 2101 は、減算器 1606 から出力された入力信号に分析窓を乗じた後、MDCT 変換(変形離散コサイン変換)して MDCT 係数を求める。MDCT 変換は、前後の隣接フレームと分析フレームを半分ずつ完全に重ね合 15 わせ、分析フレームの前半部は奇関数、後半部は偶関数という直交基底を用いる。MDCT 変換は、波形を合成する際、逆変換後の波形を重ね合わせて加算することにより、フレーム境界歪が発生しないという特徴がある。MDCT を行う際には、sin 窓などの窓関数を入力信号に乘ずる。MDCT 係数を  $X(n)$  とすると、MDCT 係数は、式 (42) に従い算出される。

$$X(m) = \sqrt{\frac{1}{N}} \sum_{n=0}^{2N-1} x(n) \cos \left\{ \frac{(2n+1+N) \cdot (2m+1)\pi}{4N} \right\}$$

… (4 2)

MDCT係数量子化器2102は、MDCT部2101から出力された入力  
 5 信号に周波数決定部1607から出力された量子化の対象となる周波数に対応  
 する係数を量子化する。そして、MDCT係数量子化器2102は、量子化し  
 たMDCT係数の符号化コードを多重化器1609に出力する。

このように、本実施の形態の音響符号化装置によれば、基本レイヤの符号化  
 10 コードを復号化した信号から拡張レイヤの符号化の対象となる周波数を決定す  
 ることにより、符号化側から復号化側に伝送する基本レイヤの符号化信号のみ  
 で拡張レイヤの符号化の対象となる周波数を決定することができ、符号化側か  
 ら復号化側にこの周波数の情報を伝送する必要がなくなり、低ビットレートで  
 高品質に符号化を行うことができる。

なお、上記実施の形態では、FFTを使った聴覚マスキングの算出法につい  
 15 て説明しているが、FFTの代わりMDCTを使って聴覚マスキングを算出す  
 ることもできる。図23は、本実施の形態の周波数決定部の内部構成の一例を  
 示すブロック図である。但し、図21と同一の構成となるものについては、図  
 21と同一番号を付し、詳しい説明を省略する。

MDCT部2201は、MDCT係数を使って振幅スペクトルP (m) を近  
 20 似する。具体的には、MDCT部2201は、以下の式(4 3)を用いてP (m)  
 を近似する。

$$P(m) = \sqrt{R^2(m)} \quad \dots (43)$$

ここで、 $R(m)$  は、アップサンプリング器 1604 から与えられる信号を M  
D C T 変換して求めた M D C T 係数を表す。

5 推定聴覚マスキング算出器 1902 は、M D C T 部 2201 において近似さ  
れた  $P(m)$  からバーカスペクトル  $B(k)$  を算出する。それ以後は上述した  
方法に従い量子化の対象となる周波数情報を算出する。

このように、本実施の形態の音響符号化装置は、M D C T を使って聴覚マス  
キングを算出することもできる。

次に、復号化側について説明する。図 24 は、本発明の実施の形態 9 に係る  
10 音響復号化装置の構成を示すブロック図である。図 24 の音響復号化装置 23  
00 は、分離器 2301 と、基本レイヤ復号化器 2302 と、アップサンプリ  
ング器 2303 と、周波数決定部 2304 と、拡張レイヤ復号化器 2305 と、  
加算器 2306 とから主に構成される。

15 分離器 2301 は、音響符号化装置 1600 において符号化されたコードを  
基本レイヤ用の第 1 符号化コードと拡張レイヤ用の第 2 符号化コードに分離し、  
第 1 符号化コードを基本レイヤ復号化器 2302 に出力し、第 2 符号化コード  
を拡張レイヤ復号化器 2305 に出力する。

20 基本レイヤ復号化器 2302 は、第 1 符号化コードを復号してサンプリング  
レート F L の復号信号を得る。そして、基本レイヤ復号化器 2302 は、復号  
信号をアップサンプリング器 2303 に出力する。アップサンプリング器 23  
03 は、サンプリングレート F L の復号信号をサンプリングレート F H の復号  
信号に変換して周波数決定部 2304 と加算器 2306 に出力する。

周波数決定部 2304 は、アップサンプリングされた基本レイヤの復号信号

を用いて拡張レイヤ復号化器 2305 で復号化の対象となる誤差スペクトルの周波数を決定する。この周波数決定部 2304 は、図 16 の周波数決定部 1607 と同様の構成をとる。

拡張レイヤ復号化器 2305 は、第 2 符号化コードを復号してサンプリング 5 レート F H の復号信号を得る。そして、拡張レイヤ復号化器 2305 は、復号された拡張フレーム単位の復号信号を重ね合わせ、重ね合わせた復号信号を加算器 2306 に出力する。具体的には、拡張レイヤ復号化器 2305 は、復号信号に合成用の窓関数を乗じ、前フレームで復号された時間領域の信号とフレームの半分だけオーバーラップさせて加算して出力信号を生成する。

10 加算器 2306 は、アップサンプリング器 2303 においてアップサンプリングされた基本レイヤの復号信号と、拡張レイヤ復号化器 2305 において復号化された拡張レイヤの復号信号とを加算して出力する。

次に、本実施の形態に係る音響復号化装置の各ブロックの詳細な動作について説明する。図 25 は、本実施の形態の音響復号化装置の拡張レイヤ復号化器 15 の内部構成の一例を示すブロック図である。図 25 は、図 24 の拡張レイヤ復号化器 2305 の内部構成の一例を示す図である。図 25 の拡張レイヤ復号化器 2305 は、MDCT 係数復号化器 2401 と、IMDCT 部 2402 と、重ね合わせ加算器 2403 とから主に構成される。

MDCT 係数復号化器 2401 は、周波数決定部 2304 から出力される復号化の対象となる誤差スペクトルの周波数に基づいて分離器 2301 から出力される第 2 符号化コードから量子化された MDCT 係数を復号する。具体的には、周波数決定部 2304 から示された信号の周波数に対応する復号 MDCT 係数を配置し、それ以外の周波数にはゼロを与える。

IMDCT 部 2402 は、MDCT 係数復号化器 2401 から出力される MDCT 係数に逆 MDCT 変換を施し、時間領域の信号を生成して重ね合わせ加算器 2403 に出力する。

重ね合わせ加算器 2403 は、復号された拡張フレーム単位の復号信号を重ね合わせ、重ね合わせた復号信号を加算器 2306 に出力する。具体的には、重ね合わせ加算器 2403 は、復号信号に合成用の窓関数を乗じ、前フレームで復号された時間領域の信号とフレームの半分だけオーバーラップさせて加算 5 して出力信号を生成する。

このように、本実施の形態の音響復号化装置によれば、基本レイヤの符号化コードを復号化した信号から拡張レイヤの復号化の対象となる周波数を決定することにより、符号化側から復号化側に伝送する基本レイヤの符号化コードのみで拡張レイヤの復号化の対象となる周波数を決定することができ、符号化側 10 から復号化側にこの周波数の情報を伝送する必要がなくなり、低ビットレートで高品質に符号化を行うことができる。

(実施の形態 10)

本実施の形態では、基本レイヤの符号化において C E L P を用いる例について説明する。図 26 は、本発明の実施の形態 10 の基本レイヤ符号化器の内部構成の一例を示すブロック図である。図 26 は、図 16 の基本レイヤ符号化器 1602 の内部構成を示す図である。図 26 の基本レイヤ符号化器 1602 は、LPC 分析器 2501 と、聴感重み部 2502 と、適応符号帳探索器 2503 と、適応ゲイン量子化器 2504 と、目標ベクトル生成器 2505 と、雑音符号帳探索器 2506 と、雑音ゲイン量子化器 2507 と、多重化器 2508 と 20 から主に構成される。

LPC 分析器 2501 は、サンプリングレート F L の入力信号の LPC 係数を算出し、この LPC 係数を LSP 係数などの量子化に適したパラメータに変換して量子化する。そして、LPC 分析器 2501 は、この量子化で得られる符号化コードを多重化器 2508 に出力する。

25 また、LPC 分析器 2501 は、符号化コードから量子化後の LSP 係数を算出して LPC 係数に変換し、量子化後の LPC 係数を、適応符号帳探索器 2

503、適応ゲイン量子化器2504、雑音符号帳探索器2506、及び雑音ゲイン量子化器2507に出力する。さらに、LPC分析器2501は、量子化前のLPC係数を聴感重み部2502、適応符号帳探索器2503、適応ゲイン量子化器2504、雑音符号帳探索器2506、及び雑音ゲイン量子化器2507に出力する。

5 2502は、LPC分析器2501で求められたLPC係数に基づいてダウンサンプリング器1601から出力された入力信号に聴感重み付けを行う。これは、量子化歪のスペクトルを入力信号のスペクトル包絡にマスクされるようスペクトル整形を行うことを目的としている。

10 適応符号帳探索器2503では、聴覚重み付けされた入力信号を目標信号として適応符号帳の探索が行われる。過去の音源系列をピッチ周期で繰り返した信号を適応ベクトルと呼び、あらかじめ定められた範囲のピッチ周期で生成された適応ベクトルによって適応符号帳は構成される。

15 聴覚重み付けされた入力信号を  $t_i(n)$  、ピッチ周期  $i$  の適応ベクトルに量子化前のLPC係数と量子化後のLPC係数で構成される重み付き合成フィルタのインパルス応答を畳み込んだ信号を  $p_i(n)$  としたとき、適応符号帳探索器2503は、式(44)の評価関数Dを最小とする適応ベクトルのピッチ周期  $i$  をパラメータとして多重化器2508に出力する。

$$D = \sum_{n=0}^{N-1} t^2(n) - \frac{\left( \sum_{n=0}^{N-1} t(n) p_i(n) \right)^2}{\sum_{n=0}^{N-1} p_i^2(n)} \quad \dots (44)$$

ここで、Nはベクトル長を表す。式(44)の第1項はピッチ周期iに独立なので、実際には、適応符号帳探索器2503は第2項のみを計算する。

5 適応ゲイン量子化器2504は、適応ベクトルに乘じられる適応ゲインの量子化を行う。適応ゲイン $\beta$ は、以下の式(45)で表され、適応ゲイン量子化器2504は、この適応ゲイン $\beta$ をスカラー量子化し、量子化時に得られる符号を多重化器2508に出力する。

$$\beta = \frac{\sum_{n=0}^{N-1} t(n) p_i(n)}{\sum_{n=0}^{N-1} p_i^2(n)} \quad \dots (45)$$

目標ベクトル生成器 2505 は、入力信号から適応ベクトルの影響を減算して、雑音符号帳探索器 2506 と雑音ゲイン量子化器 2507 で用いる目標ベクトルを生成して出力する。目標ベクトル生成器 2505 は、 $p_i(n)$  を式 12 で表される評価関数 D を最小とするときの適応ベクトルに重み付き合成フィルタのインパルス応答を畠み込んだ信号、 $\beta q$  を式 13 で表される適応ベクトル  $\beta$  をスカラー量子化したときの量子化値としたとき、目標ベクトル  $t_2(n)$  は、以下に示す式 (46) のように表される。

10

$$t_2(n) = t(n) - \beta q \cdot p_i(n) \quad \dots (46)$$

雑音符号帳探索器 2506 は、前記目標ベクトル  $t_2(n)$  と量子化前の L P C 係数と量子化後の L P C 係数を用いて雑音符号帳の探索を行う。例えば、  
15 雜音符号帳探索器 2506 には、ランダム雑音や大規模な音声信号を使って学

習した信号を用いることができる。また、雑音符号帳探索器 2506 が備える雑音符号帳は、代数(Algebraic)符号帳のように、振幅 1 のパルスをあらかじめ定められた非常に少ない数だけ有するベクトルで表されることができる。この代数符号長は、パルスの位置とパルスの符号(極性)の最適な組み合わせを少な

5 い計算量で決定することができるという特徴がある。

雑音符号帳探索器 2506 は、目標ベクトルを  $t_2(n)$ 、コード  $j$  に対応する雑音ベクトルに重み付き合成フィルタのインパルス応答を畳み込んだ信号を  $c_j(n)$  としたとき、以下に示す式 (47) の評価関数  $D$  を最小とする雑音ベクトルのインデックス  $j$  を多重化器 2508 に出力する。

10

$$D = \sum_{n=0}^{N-1} t_2^2(n) - \frac{\left( \sum_{n=0}^{N-1} t_2(n) c_j(n) \right)^2}{\sum_{n=0}^{N-1} c_j^2(n)} \quad \dots (47)$$

雑音ゲイン量子化器 2507 は、雑音ベクトルに乘じる雑音ゲインを量子化する。雑音ゲイン量子化器 2507 は、以下に示す式 (48) を用いて雑音ゲイン  $\gamma$  を算出し、この雑音ゲイン  $\gamma$  をスカラー量子化して多重化器 2508 に

15 出力する。

$$\gamma = \frac{\sum_{n=0}^{N-1} t_2(n) c_j(n)}{\sum_{n=0}^{N-1} c_j^2(n)} \quad \dots (48)$$

多重化器 2508 は、送られてきた L P C 係数、適応ベクトル、適応ゲイン、  
5 雜音ベクトル、雜音ゲインの符号化コードを多重化して局所復号化器 1603  
及び多重化器 1609 に出力する。

次に、復号化側について説明する。図 27 は、本実施の形態の基本レイヤ復  
号化器の内部構成の一例を示すブロック図である。図 27 は、図 24 の基本レ  
イヤ復号化器 2302 の内部構成を示す図である。図 27 の基本レイヤ復号化  
器 2302 は、分離器 2601 と、音源生成器 2602 と、合成フィルタ 26  
10 03 とから主に構成される。

分離器 2601 は、分離器 2301 から出力された第 1 符号化コードを L P  
C 係数、適応ベクトル、適応ゲイン、雜音ベクトル、雜音ゲインの符号化コー  
ドに分離して、適応ベクトル、適応ゲイン、雜音ベクトル、雜音ゲインの符号  
化コードを音源生成器 2602 に出力する。同様に、分離器 2601 は、 L P  
15 C 係数の符号化コードを合成フィルタ 2603 に出力する。

音源生成器 2602 は、適応ベクトル、適応ベクトルゲイン、雜音ベクトル、  
雜音ベクトルゲインの符号化コードを復号し、以下に示す式 (49) を用いて  
音源ベクトル  $e_x(n)$  を生成する。

$$ex(n) = \beta_q \cdot q(n) - \gamma_q \cdot c(n) \quad \dots (49)$$

ここで、 $q(n)$  は適応ベクトル、 $\beta_q$  は適応ベクトルゲイン、 $c(n)$  は雑音ベクトル、 $\gamma_q$  は雑音ベクトルゲインを表す。

5 合成フィルタ 2603 では、LPC 係数の符号化コードから LPC 係数を復号し、以下に示す式 (50) を用いて復号された LPC 係数から合成信号  $syn(n)$  を生成する。

$$syn(n) = ex(n) + \sum_{i=1}^{NP} \alpha_q(i) \cdot syn(n-i) \quad \dots (50)$$

10

ここで、 $\alpha_q$  は復号された LPC 係数、NP は LPC 係数の次数を表す。そして、合成フィルタ 2603 は、復号された復号信号  $syn(n)$  をアップサンプリング器 2303 に出力する。

15 このように、本実施の形態の音響符号化装置及び音響復号化装置によれば、送信側において、基本レイヤに CELP を適用して入力信号を符号化し、受信側において、この符号化した入力信号に CELP を適用して復号することにより、低ビットレートで高品質な基本レイヤを実現することができる。

なお、本実施の形態の音声符号化装置は、量子化歪の知覚を抑制するために、合成フィルタ 2603 の後にポストフィルタを従属接続する構成を採ることもできる。図 28 は、本実施の形態の基本レイヤ復号化器の内部構成の一例を示すブロック図である。但し、図 27 と同一の構成となるものについては、図 27 と同一番号を付し、詳しい説明を省略する。

ポストフィルタ 2701 は、量子化歪の知覚の抑制の実現のために様々な構成を適用しうるが、代表的な方法として、分離器 2601 で復号されて得られる L P C 係数から構成されるホルマント強調フィルタを用いる方法がある。ホルマント強調フィルタ  $H_f(z)$  は以下に示す式 (51) で表される。

5

$$H_f(z) = \frac{A(z/\gamma_n)}{A(z/\gamma_d)} (1 - \mu z^{-1})$$

… (51)

ここで、 $A(z)$  は復号 L P C 係数から構成される合成フィルタ、 $\gamma_n$ 、 $\gamma_d$ 、  
10  $\mu$  はフィルタの特性を決定する定数を表す。

(実施の形態 1 1)

図 29 は、本発明の実施の形態 1 1 に係る音響符号化装置の周波数決定部の内部構成の一例を示すブロック図である。但し、図 20 と同一の構成となるものについては、図 20 と同一番号を付し、詳しい説明を省略する。図 29 の周波数決定部 1607 は、推定誤差スペクトル算出器 2801 と、決定部 2802 とを具備し、基本レイヤ復号信号の振幅スペクトル  $P(m)$  から推定誤差スペクトル  $E'(m)$  を推定し、推定誤差スペクトル  $E'(m)$  と推定聴覚マスキング  $M'(m)$  とを用いて拡張レイヤ符号化器 1608 で符号化される誤差スペクトルの周波数を決定する点が図 20 と異なる。

20 F F T 部 1901 は、アップサンプリング器 1604 から出力された基本レイヤ復号信号  $x(n)$  を直交変換して振幅スペクトル  $P(m)$  を算出して推定聴覚マスキング算出器 1902 と推定誤差スペクトル算出器 2801 に出力す

る。

推定誤差スペクトル算出器 2801 は、FFT部 1901 で算出される基本レイヤ復号信号の振幅スペクトル  $P(m)$  から推定誤差スペクトル  $E'(m)$  を算出して決定部 2802 に出力する。推定誤差スペクトル  $E'(m)$  は、基 5 本レイヤ復号信号の振幅スペクトル  $P(m)$  を平坦に近づける処理を施し算出される。具体的には、推定誤差スペクトル算出器 2801 は、以下の式 (52) を用いて推定誤差スペクトル  $E'(m)$  を算出する。

$$E'(m) = a \cdot P(m)^\gamma \quad \dots (52)$$

10

ここで  $a$  と  $\gamma$  は 0 以上 1 未満の定数を表す。

決定部 2802 は、推定誤差スペクトル算出器 2801 において推定された推定誤差スペクトル  $E'(m)$  と推定聴覚マスキング算出器 1902 で得られる推定聴覚マスキング  $M'(m)$  を用いて拡張レイヤ符号化器 1608 で誤差 15 スペクトルを符号化する対象の周波数を決定する。

次に、本実施の形態の推定誤差スペクトル算出器 2801 が算出する推定誤差スペクトルについて説明する。図 30 は、本実施の形態の推定誤差スペクトル算出器が算出する残差スペクトルの一例を示す図である。

誤差スペクトル  $E(m)$  は、図 30 に示すように基本レイヤ復号信号の振幅 20 スペクトル  $P(m)$  に比べスペクトルの形状が平坦になり、かつ全帯域のパワーが小さくなっている。よって、振幅スペクトル  $P(m)$  を  $\gamma$  ( $0 < \gamma < 1$ ) 乗することによりスペクトル形状を平坦化し、 $a$  ( $0 < a < 1$ ) 倍することにより全体域のパワーを減少させることにより、誤差スペクトルの推定精度を向

上させることができる。

同様に、復号化側も音響復号化装置 2300 の周波数決定部 2304 の内部構成を符号化側の図 29 の周波数決定部 1607 と同じ構成とする。

5 このように、本実施の形態の音響符号化装置によれば、基本レイヤの復号信号のスペクトルから推定した残差スペクトルを平滑化することにより、推定誤差スペクトルを残差スペクトルに近似することができ、拡張レイヤにて誤差スペクトルを効率よく符号化することができる。

なお、本実施の形態では、FFT を用いた場合について説明したが、前述した実施の形態 9 と同様に、FFT の代わりに MDCT を用いる構成も可能である。

#### (実施の形態 12)

図 31 は、本発明の実施の形態 12 に係る音響符号化装置の周波数決定部の内部構成の一例を示すブロック図である。但し、図 20 と同一の構成となるものについては、図 20 と同一番号を付し、詳しい説明を省略する。図 31 の周波数決定部 1607 は、推定聴覚マスキング修正部 3001 と、決定部 3002 を具備し、周波数決定部 1607 において、基本レイヤ復号信号の振幅スペクトル  $P$  (m) から推定聴覚マスキング算出器 1902 にて推定聴覚マスキング  $M'$  (m) を算出した後に、この推定聴覚マスキング  $M'$  (m) に局所復号化器 1603 の復号パラメータの情報を基に修正を加える点が図 20 と異なる。

FFT 部 1901 は、アップサンプリング器 1604 から出力された基本レイヤ復号信号  $x$  (n) を直交変換して振幅スペクトル  $P$  (m) を算出して推定聴覚マスキング算出器 1902 と決定部 3002 に出力する。推定聴覚マスキング算出器 1902 は、基本レイヤ復号信号の振幅スペクトル  $P$  (m) を用いて推定聴覚マスキング  $M'$  (m) を算出して推定聴覚マスキング修正部 3001 に出力する。

推定聴覚マスキング修正部 3001 は、局所復号化器 1603 から入力される基本レイヤの復号パラメータの情報を用いて推定聴覚マスキング算出器 1902 で求められる推定聴覚マスキング  $M'$  (m) に修正を加える。

ここでは、基本レイヤの符号化コードの情報として、復号 L P C 係数から算出される 1 次の PARCOR 係数が与えられるものとする。一般に L P C 係数や PARCOR 係数は入力信号のスペクトル包絡を表現する。PARCOR 係数の次数を下げていくと、PARCOR 係数の性質から、スペクトル包絡の形状が簡略化されてゆき、PARCOR 係数の次数が 1 次のときにスペクトルの傾きの程度を表すようになる。

一方で、入力信号として与えられる楽音や音声のスペクトル特性には、高域に対して低域にパワーが偏っている場合（例えば母音）やその逆の場合（例えば子音）が存在する。基本レイヤ復号信号はこういった入力信号のスペクトル特性に影響を受けやすく、必要以上にスペクトルのパワーの偏りを強調してしまう傾向にある。

そこで、本実施の形態の音響符号化装置は、推定聴覚マスキング修正部 3001 において、前述した 1 次の PARCOR 係数を利用して過度に強調されたスペクトルの偏りを補正することにより、推定マスキング  $M'$  (m) の精度を向上させることができる。

推定聴覚マスキング修正部 3001 は、以下に示す式 (53) を用いて基本レイヤ符号化器 1602 から出力された 1 次の PARCOR 係数  $k(1)$  から修正フィルタ  $H_k(z)$  を算出する。

$$H_k(z) = 1 - \beta \cdot k(1) \cdot z^{-1} \quad \dots (53)$$

ここで  $\beta$  は 1 未満の正の定数を表す。次に、推定聴覚マスキング修正部 300  
1 は、以下に示す式 (54) を用いて  $H_k(z)$  の振幅特性  $K(m)$  を算出す  
る。

5

$$K(m) = \left| 1 - \beta \cdot k(1) \cdot e^{-j \frac{2\pi m}{M}} \right| \quad \dots (54)$$

そして、推定聴覚マスキング修正部 300 1 は、以下の式 (55) を用いて  
修正フィルタの振幅特性  $K(m)$  から修正後の推定聴覚マスキング  $M'(m)$   
を算出する。

10

$$M''(m) = K(m) \cdot M'(m) \quad \dots (55)$$

15

そして、推定聴覚マスキング修正部 300 1 は、推定聴覚マスキング  $M'(m)$   
の代わりに修正後の聴覚マスキング  $M''(m)$  を決定部 300 2 に出力する。

決定部 300 2 は、基本レイヤ復号信号の振幅スペクトル  $P(m)$  と推定聴  
覚マスキング修正部 300 1 から出力される修正後の聴覚マスキング  $M''(m)$

(m) を用いて拡張レイヤ符号化器 1608 で誤差スペクトルを符号化する対象の周波数を決定する。

5 このように本実施の形態の音響符号化装置によれば、マスキング効果の特性を利用して、入力信号のスペクトルから聴覚マスキングを算出し、拡張レイヤの符号化において、量子化歪をこのマスキング値以下になるように量子化を行うことにより、品質の劣化を伴わずに量子化の対象となるMDCT係数の数を減らすことができ、低ビットレートで高品質に符号化を行うことができる。

10 このように、本実施の形態の音響符号化装置によれば、基本レイヤ復号信号の振幅スペクトルから推定した推定聴覚マスキングを、基本レイヤ符号化器の復号パラメータの情報を基に修正を加えることにより、推定聴覚マスキングの精度を向上させることができ、結果拡張レイヤにて誤差スペクトルを効率よく符号化することができる。

同様に、復号化側も音響復号化装置 2300 の周波数決定部 2304 の内部構成を符号化側の図 31 の周波数決定部 1607 と同じ構成とする。

15 なお、本実施の形態の周波数決定部 1607 は、本実施の形態と実施の形態 11 を組み合わせた構成を探ることもできる。図 32 は、本実施の形態の音響符号化装置の周波数決定部の内部構成の一例を示すブロック図である。但し、図 20 と同一の構成となるものについては、図 20 と同一番号を付し、詳しい説明を省略する。

20 FFT 部 1901 は、アップサンプリング器 1604 から出力された基本レイヤ復号信号  $x(n)$  を直交変換して振幅スペクトル  $P(m)$  を算出して推定聴覚マスキング算出器 1902 と推定誤差スペクトル算出器 2801 に出力する。

25 推定聴覚マスキング算出器 1902 は、基本レイヤ復号信号の振幅スペクトル  $P(m)$  を用いて推定聴覚マスキング  $M'(m)$  を算出して推定聴覚マスキング修正部 3001 に出力する。

推定聴覚マスキング修正部 3001 は、局所復号化器 1603 から入力される基本レイヤの復号パラメータの情報が推定聴覚マスキング修正部 3001 を用いて推定聴覚マスキング算出器 1902 で求められる推定聴覚マスキング  $M'$  (m) に修正を加える。

5 推定誤差スペクトル算出器 2801 は、FFT部 1901 で算出される基本レイヤ復号信号の振幅スペクトル  $P$  (m) から推定誤差スペクトル  $E'$  (m) を算出して決定部 3101 に出力する。

10 決定部 3101 は、推定誤差スペクトル算出器 2801 において推定された推定誤差スペクトル  $E'$  (m) と推定聴覚マスキング修正部 3001 から出力される修正後の聴覚マスキング  $M'$  (m) を用いて拡張レイヤ符号化器 1608 で誤差スペクトルを符号化する対象の周波数を決定する。

また、本実施の形態では、FFTを用いた場合について説明したが、前述した実施の形態 9 と同様に、FFTの代わりにMDCTを用いる構成も可能である。

15 (実施の形態 13)

図 33 は、本発明の実施の形態 13 に係る音響符号化装置の拡張レイヤ符号化器の内部構成の一例を示すブロック図である。但し、図 22 と同一の構成となるものについては、図 22 と同一番号を付し、詳しい説明を省略する。図 33 の拡張レイヤ符号化器は、順序づけ部 3201 と、MDCT係数量子化器 3202 を具備し、周波数決定部 1607 から与えられる周波数を推定歪値  $D$  (m) の大きさに従い周波数別で符号化後の情報量に重み付けを行う点が図 22 の拡張レイヤ符号化器と異なる。

20 図 33において、MDCT部 2101 は、減算器 1606 から出力された入力信号に分析窓を乗じた後、MDCT変換(変形離散コサイン変換)してMDCT係数を求め、MDCT係数量子化器 3202 に出力する。

25 順序づけ部 3201 は、周波数決定部 1607 で求められた周波数情報を受

けつけ、各周波数の推定誤差スペクトル  $E'$  (m) が推定聴覚マスキング  $M'$  (m) を超える量 (以後、推定歪値と呼ぶ)  $D$  (m) を算出する。この推定歪値  $D$  (m) は、以下に示す式 (5 6) で定義される。

$$5 \quad D \text{ (m)} = E' \text{ (m)} - M' \text{ (m)} \quad \cdots (5 6)$$

ここで、順序づけ部 3201 は、以下に示す式 (5 7) を満たす推定歪値  $D$  (m) のみ算出する。

$$10 \quad E' \text{ (m)} - M' \text{ (m)} > 0 \quad \cdots (5 7)$$

そして、順序づけ部 3201 は、推定歪値  $D$  (m) の大きさが大きいものから順序付けを行い、その周波数情報を MDC T 係数量子化器 3202 に出力する。MDC T 係数量子化器 3202 では、推定歪値  $D$  (m) により順序付けされた周波数情報を基に、推定歪値  $D$  (m) の大きいものからその周波数に位置する誤差スペクトル  $E$  (m) にビットを多く配分して量子化を行う。

ここでは例として、周波数決定手段から送られてくる周波数と推定歪値が図 34 である場合について説明する。図 34 は、本実施の形態の順序づけ部の推定歪値の順位づけの一例を示す図である。

20 順序づけ部 3201 は、図 34 の情報から、推定歪値  $D$  (m) が大きい順に周波数の並べ替えを行う。この例では、順序づけ部 3201 の処理の結果、周波数  $m = 7, 8, 4, 9, 1, 11, 3, 12$  の順序が得られる。順序づけ部

3201は、この順序付けの情報をMDCT係数量子化器3202に出力する。

MDCT係数量子化器3202では、MDCT部2101から与えられる誤差スペクトルE (m) の内、順序づけ部3201から与えられる順序付けの情報を基にして、E (7)、E (8)、E (4)、E (9)、E (1)、E (15)、E (3)、E (12)を量子化する。

このとき、順序づけの先頭に位置する誤差スペクトルの量子化に用いられるビット数は多く配分され、末尾にいくに従いビット数は少なく配分される。すなわち、推定歪値D (m) が大きい周波数ほど誤差スペクトルの量子化に用いられるビット数は多く配分され、推定歪値D (m) が小さい周波数ほど誤差スペクトルの量子化に用いられるビット数は少なく配分される。

例えば、E (7)を8ビット、E (8)、E (4)を7ビット、E (9)、E (1)を6ビット、E (11)、E (3)、E (12)を5ビットというようなビット配分を実施する。このような推定歪値D (m) に応じた適応ビット配分を行うことにより、量子化の効率が向上することになる。

ベクトル量子化を適用する場合には、拡張レイヤ符号化器1608は、先頭に位置する誤差スペクトルから順にベクトルを構成し、それぞれのベクトルに對してベクトル量子化を行う。このとき、先頭に位置する誤差スペクトルのビット配分が多くなり、末尾に位置する誤差スペクトルのビット配分が小さくなるようなベクトルの構成と量子化ビット配分が成される。図34の例では、V1 = (E (7)、E (8))、V2 = (E (4)、E (9))、V3 = (E (1)、E (11)、E (3)、E (12))のように、2次元、2次元、4次元の3ベクトルを構成し、V1を10ビット、V2を8ビット、V3を8ビットというようなビット配分を行う。

このように、本実施の形態の音響符号化装置によれば、拡張レイヤでの符号化において、推定誤差スペクトルが推定聴覚マスキングを超える量が大きい周波数に多くの情報量を配分して符号化することにより、量子化効率の向上を図

ることができる。

次に復号化側について説明する。図35は、本発明の実施の形態13に係る音響復号化装置の拡張レイヤ復号化器の内部構成の一例を示すブロック図である。但し、図25と同一の構成となるものについては、図25と同一番号を付5し、詳しい説明を省略する。図35の拡張レイヤ復号化器2305は、順序づけ部3401と、MDCT係数復号化器3402とを具備し、周波数決定部2304から与えられる周波数を推定歪値D (m) の大きさに従い順序付けを行う点が図25と異なる。

順序づけ部3401は、上述の式(56)を用いて推定歪値D (m) を算出10する。順序づけ部3401は、上述の順序づけ部3201と同一の構成を採る。この構成により適応ビット配分を行い量子化効率の向上を図ることができる上述の音響符号化法の符号化コードを復号することができる。

MDCT係数復号化器3402は、推定歪値D (m) の大きさに従い順序付けされた周波数の情報を用いて分離器2301から出力された第2符号化コードを復号化する。具体的には、MDCT係数復号化器3402は、周波数決定部2304から与えられる周波数に対応する復号MDCT係数を配置し、それ以外の周波数にはゼロを与える。次にIMDCT部2402は、MDCT係数復号化器3402から得られるMDCT係数に逆MDCT変換を施し、時間領域の信号を生成する。

重ね合わせ加算器2403は、前記信号に合成用の窓関数を乗じ、前フレームで復号された時間領域の信号とフレームの半分だけオーバーラップさせて加算して出力信号を生成する。重ね合わせ加算器2403は、この出力信号を加算器2306に出力する。

このように、本実施の形態の音響復号化装置によれば、拡張レイヤでの符号化において、推定誤差スペクトルが推定聴覚マスキングを超える量に応じて適応ビット配分されたベクトル量子化を行うことにより、量子化効率の向上を図

ることができる。

(実施の形態 14)

図 36 は、本発明の実施の形態 14 に係る音響符号化装置の拡張レイヤ符号化器の内部構成の一例を示すブロック図である。但し、図 22 と同一の構成と 5 なるものについては、図 22 と同一番号を付し、詳しい説明を省略する。図 3 6 の拡張レイヤ符号化器は、固定帯域指定部 3501 と、MDCT 係数量子化器 3502 とを具備し、あらかじめ定めておいた帯域に含まれる MDCT 係数を周波数決定部 1607 から得られる周波数と共に量子化する点が図 22 の拡張レイヤ符号化器と異なる。

10 図 36において、固定帯域指定部 3501 には、あらかじめ聴感上重要な帯域が設定されている。ここでは、設定されている帯域に含まれる周波数を  $m = 15, 16$  とする。

MDCT 係数量子化器 3502 は、MDCT 部 2101 から出力された入力信号に周波数決定部 1607 から出力された聴覚マスキングを用いて入力信号 15 を量子化する係数と量子化しない係数に分類し、量子化する係数と、さらに固定帯域指定部 3501 が設定する帯域にある係数を符号化する。

その周波数が図 34 で示されたものであるとすると、MDCT 係数量子化器 3502 では、誤差スペクトル  $E(1), E(3), E(4), E(7), E(8), E(9), E(11), E(12)$  および、固定帯域指定部 3501 20 で指定される周波数の誤差スペクトル  $E(15), E(16)$  が量子化される。

このように、本実施の形態の音響符号化装置によれば、符号化の対象として選択されにくいが聴覚的に重要な帯域を強制的に量子化することにより、本来符号化の対象として選択されるべき周波数が選択されない場合でも、聴覚的に重要な帯域に含まれる周波数に位置する誤差スペクトルは必ず量子化されるこ 25 とになり、品質を改善することができる。

次に、復号化側について説明する。図 37 は、本発明の実施の形態 14 に係

る音響復号化装置の拡張レイヤ復号化器の内部構成の一例を示すブロック図である。但し、図25と同一の構成となるものについては、図25と同一番号を付し、詳しい説明を省略する。図37の拡張レイヤ復号化器は、固定帯域指定部3601と、MDCT係数復号化器3602とを具備し、あらかじめ定めて  
5 おいた帯域に含まれるMDCT係数を周波数決定部2304から得られる周波数と共に復号化する点が図25の拡張レイヤ復号化器と異なる。

図37において、固定帯域指定部3601には、あらかじめ聴感上重要な帯域が設定されている。

MDCT係数復号化器3602は、周波数決定部2304から出力される復号化の対象となる誤差スペクトルの周波数に基づいて分離器2301から出力される第2符号化コードから量子化されたMDCT係数を復号する。具体的には、周波数決定部2304と固定帯域指定部3601から示された周波数に対応する復号MDCT係数を配置し、それ以外の周波数にはゼロを与える。

IMDCT部2402は、MDCT係数復号化器3602から出力されるMDCT係数に逆MDCT変換を施し、時間領域の信号を生成して重ね合わせ加算器2403に出力する。

このように、本実施の形態の音響復号化装置によれば、あらかじめ定めておいた帯域に含まれるMDCT係数を復号化することにより、符号化の対象として選択されにくいが聴覚的に重要な帯域を強制的に量子化された信号を復号化  
20 することができ、符号化側において本来符号化の対象として選択されるべき周波数が選択されない場合でも、聴覚的に重要な帯域に含まれる周波数に位置する誤差スペクトルは必ず量子化されることになり、品質を改善させることができる。

なお、本実施の形態の拡張レイヤ符号化器及び拡張レイヤ復号化器は、本実  
25 施の形態と実施の形態13とを組み合わせた構成を探ることもできる。図38  
は、本実施の形態の音響符号化装置の周波数決定部の内部構成の一例を示すブ

ロック図である。但し、図22と同一の構成となるものについては、図22と同一番号を付し、詳しい説明を省略する。

図38において、MDCT部2101は、減算器1606から出力された入力信号に分析窓を乗じた後、MDCT変換(変形離散コサイン変換)してMDCT係数を求め、MDCT係数量子化器3701に出力する。

順序づけ部3201は、周波数決定部1607で求められた周波数情報を受けつけ、各周波数の推定誤差スペクトル $E'$  (m) が推定聴覚マスキング $M'$  (m) を超える量(以後、推定歪値と呼ぶ)  $D$  (m) を算出する。

固定帯域指定部3501には、あらかじめ聴感上重要な帯域が設定されている。

MDCT係数量子化器3701では、推定歪値 $D$  (m) により順序付けされた周波数情報を基に、推定歪値 $D$  (m) の大きいものからその周波数に位置する誤差スペクトル $E$  (m) にビットを多く配分して量子化を行う。また、MDCT係数量子化器3701は、固定帯域指定部3501が設定する帯域にある係数を符号化する。

次に復号化側について説明する。図39は、本発明の実施の形態14に係る音響復号化装置の拡張レイヤ復号化器の内部構成の一例を示すブロック図である。但し、図25と同一の構成となるものについては、図25と同一番号を付し、詳しい説明を省略する。

図39において、順序づけ部3401は、周波数決定部2304で求められた周波数情報を受けつけ、各周波数の推定誤差スペクトル $E'$  (m) が推定聴覚マスキング $M'$  (m) を超える量(以後、推定歪値と呼ぶ)  $D$  (m) を算出する。

そして、順序づけ部3401は、推定歪値 $D$  (m) の大きさが大きいものから順序付けを行い、その周波数情報をMDCT係數復号化器3801に出力する。固定帯域指定部3601には、あらかじめ聴感上重要な帯域が設定されている。

MDCT係数復号化器3801は、順序づけ部3401から出力される復号化の対象となる誤差スペクトルの周波数に基づいて分離器2301から出力される第2符号化コードから量子化されたMDCT係数を復号する。具体的には、順序づけ部3401と固定帯域指定部3601から示された信号の周波数に対応する復号MDCT係数を配置し、それ以外の周波数にはゼロを与える。

IMDCT部2402は、MDCT係数復号化器3801から出力されるMDCT係数に逆MDCT変換を施し、時間領域の信号を生成して重ね合わせ計算器2403に出力する。

(実施の形態15)

次に、本発明の実施の形態15について、図面を参照して説明する。図40は、本発明の実施の形態15に係る通信装置の構成を示すブロック図である。図40における信号処理装置3903は前述した実施の形態1から実施の形態14に示した音響符号化装置の中の1つによって構成されている点に本実施の形態の特徴がある。

図40に示すように、本発明の実施の形態15に係る通信装置3900は、入力装置3901、A/D変換装置3902及びネットワーク3904に接続されている信号処理装置3903を具備している。

A/D変換装置3902は、入力装置3901の出力端子に接続されている。信号処理装置3903の入力端子は、A/D変換装置3902の出力端子に接続されている。信号処理装置3903の出力端子はネットワーク3904に接続されている。

入力装置3901は、人間の耳に聞こえる音波を電気的信号であるアナログ信号に変換してA/D変換装置3902に与える。A/D変換装置3902はアナログ信号をデジタル信号に変換して信号処理装置3903に与える。信号処理装置3903は入力されてくるデジタル信号を符号化してコードを生成し、ネットワーク3904に出力する。

このように、本発明の実施の形態の通信装置によれば、通信において前述した実施の形態1～14に示したような効果を享受でき、少ないビット数で効率よく音響信号を符号化する音響符号化装置を提供することができる。

(実施の形態16)

5 次に、本発明の実施の形態16について、図面を参照して説明する。図41は、本発明の実施の形態16に係る通信装置の構成を示すブロック図である。図41における信号処理装置4003は前述した実施の形態1から実施の形態14に示した音響復号化装置の中の1つによって構成されている点に本実施の形態の特徴がある。

10 図41に示すように、本発明の実施の形態16に係る通信装置4000は、ネットワーク4001に接続されている受信装置4002、信号処理装置4003、及びD/A変換装置4004及び出力装置4005を具備している。

受信装置4002の入力端子は、ネットワーク4001に接続されている。信号処理装置4003の入力端子は、受信装置4002の出力端子に接続され15ている。D/A変換装置4004の入力端子は、信号処理装置4003の出力端子に接続されている。出力装置4005の入力端子は、D/A変換装置4004の出力端子に接続されている。

受信装置4002は、ネットワーク4001からのデジタルの符号化音響信号を受けてデジタルの受信音響信号を生成して信号処理装置4003に与える。信号処理装置4003は、受信装置4002からの受信音響信号を受けてこの受信音響信号に復号化処理を行ってデジタルの復号化音響信号を生成してD/A変換装置4004に与える。D/A変換装置4004は、信号処理装置4003からのデジタルの復号化音声信号を変換してアナログの復号化音声信号を生成して出力装置4005に与える。出力装置4005は、電気的25信号であるアナログの復号化音響信号を空気の振動に変換して音波として人間の耳に聴こえるように出力する。

このように、本実施の形態の通信装置によれば、通信において前述した実施の形態1～14に示したような効果を享受でき、少ないビット数で効率よく符号化された音響信号を復号することができるので、良好な音響信号を出力することができる。

5 (実施の形態17)

次に、本発明の実施の形態17について、図面を参照して説明する。図42は、本発明の実施の形態17に係る通信装置の構成を示すブロック図である。本発明の実施の形態17において、図42における信号処理装置4103は、前述した実施の形態1から実施の形態14に示した音響符号化器の中の1つに10 よって構成されている点に本実施の形態の特徴がある。

図42に示すように、本発明の実施の形態17に係る通信装置4100は、入力装置4101、A/D変換装置4102、信号処理装置4103、RF変調装置4104及びアンテナ4105を具備している。

入力装置4101は人間の耳に聞こえる音波を電気的信号であるアナログ信号に変換してA/D変換装置4102に与える。A/D変換装置4102はアナログ信号をデジタル信号に変換して信号処理装置4103に与える。信号処理装置4103は入力されてくるデジタル信号を符号化して符号化音響信号を生成し、RF変調装置4104に与える。RF変調装置4104は、符号化音響信号を変調して変調符号化音響信号を生成し、アンテナ4105に与える。アンテナ4105は、変調符号化音響信号を電波として送信する。

このように、本実施の形態の通信装置によれば、無線通信において前述した実施の形態1～14に示したような効果を享受でき、少ないビット数で効率よく音響信号を符号化することができる。

なお、本発明は、オーディオ信号を用いる送信装置、送信符号化装置又は音響信号符号化装置に適用することができる。また、本発明は、移動局装置又は基地局装置にも適用することができる。

## (実施の形態 18)

次に、本発明の実施の形態 18について、図面を参照して説明する。図 4 3 は、本発明の実施の形態 18 に係る通信装置の構成を示すブロック図である。本発明の実施の形態 18 において、図 4 3 における信号処理装置 4203 は、  
5 前述した実施の形態 1 から実施の形態 14 に示した音響復号化器の中の 1 つに  
よって構成されている点に本実施の形態の特徴がある。

図 4 3 に示すように、本発明の実施の形態 18 に係る通信装置 4200 は、  
アンテナ 4201、RF 復調装置 4202、信号処理装置 4203、D/A 変  
換装置 4204 及び出力装置 4205 を具備している。

10 アンテナ 4201 は、電波としてのデジタルの符号化音響信号を受けて電  
気信号のデジタルの受信符号化音響信号を生成して RF 復調装置 4202 に  
与える。RF 復調装置 4202 は、アンテナ 4201 からの受信符号化音響信  
号を復調して復調符号化音響信号を生成して信号処理装置 4203 に与える。

15 信号処理装置 4203 は、RF 復調装置 4202 からのデジタルの復調符  
号化音響信号を受けて復号化処理を行ってデジタルの復号化音響信号を生成  
して D/A 変換装置 4204 に与える。D/A 変換装置 4204 は、信号処理  
装置 4203 からのデジタルの復号化音声信号を変換してアナログの復号化  
音声信号を生成して出力装置 4205 に与える。出力装置 4205 は、電気的  
20 信号であるアナログの復号化音声信号を空気の振動に変換して音波として人間  
の耳に聴こえるように出力する。

このように、本実施の形態の通信装置によれば、無線通信において前述した  
実施の形態 1 ~ 14 に示したような効果を享受でき、少ないビット数で効率よ  
く符号化された音響信号を復号することができるので、良好な音響信号を出力  
することができる。

25 なお、本発明は、オーディオ信号を用いる受信装置、受信復号化装置又は音  
声信号復号化装置に適用することができる。また、本発明は、移動局装置又は

基地局装置にも適用することができる。

また、本発明は上記実施の形態に限定されず、種々変更して実施することができる。例えば、上記実施の形態では、信号処理装置として行う場合について説明しているが、これに限られるものではなく、この信号処理方法をソフトウェアとして行うことも可能である。

例えば、上記信号処理方法を実行するプログラムを予めROM (Read Only Memory) に格納しておき、そのプログラムをCPU (Central Processor Unit) によって動作させるようにしても良い。

また、上記信号処理方法を実行するプログラムをコンピュータで読み取り可能な記憶媒体に格納し、記憶媒体に格納されたプログラムをコンピュータのRAM (Random Access memory) に記録して、コンピュータをそのプログラムにしたがって動作させるようにしても良い。

なお、上記説明では、時間領域から周波数領域への変換法にMDCTを用いる場合について説明を行っているがこれに限定されず直交変換であればいずれも適用できる。例えば、離散フーリエ変換または離散コサイン変換等を適用することもできる。

なお、本発明は、オーディオ信号を用いる受信装置、受信復号化装置又は音声信号復号化装置に適用することができる。また、本発明は、移動局装置又は基地局装置にも適用することができる。

20

以上の説明から明らかなように、本発明の符号化装置、復号化装置及び符号化方法、復号化方法によれば、基本レイヤの符号化コードより得られる情報を利用して拡張レイヤの符号化を行うことにより、音声が主体で背景に音楽や雑音が重畠しているような信号であっても低ビットレートで高品質に符号化を行うことができる。

本明細書は、2002年4月26日出願の特願2002-127541及び

2002年9月12日出願の特願2002-267436に基づくものである。  
この内容をここに含めておく。

#### 産業上の利用可能性

5 本発明は、音声信号を符号化及び復号化する装置、及び通信装置に用いて好  
適である。

## 請 求 の 範 囲

1. 入力信号のサンプリングレートを下げるダウンサンプリング手段と、サンプリングレートを下げた入力信号を符号化して第1符号化コードを得る基本レイヤ符号化手段と、前記第1符号化コードに基づいて復号信号を生成する復号化手段と、前記復号信号のサンプリングレートを前記入力信号と同一のレートに上げるアップサンプリング手段と、前記復号化手段の復号処理の中で生成されたパラメータを利用し、前記入力信号と前記サンプリングレートを上げた復号信号との差分値を符号化して第2符号化コードを得る拡張レイヤ符号化手段と、前記第1符号化コードと前記第2符号化コードとを多重する多重手段とを具備する符号化装置。  
10
2. 前記基本レイヤ符号化手段は、符号励振線形予測法を用いて入力信号を符号化する請求の範囲第1項に記載の符号化装置。
3. 前記拡張レイヤ符号化手段は、直交変換を用いて入力信号を符号化する請求の範囲第1項に記載の符号化装置。
- 15 4. 前記拡張レイヤ符号化手段は、MDCT変換を用いて入力信号を符号化する請求の範囲第3項に記載の符号化装置。
5. 前記拡張レイヤ符号化手段は、前記復号化手段の復号処理の中で生成された基本レイヤのLPC係数を利用して符号化処理を行うことを特徴とする請求の範囲1から請求の範囲第4項に記載の符号化装置。
- 20 6. 前記拡張レイヤ符号化手段は、予め設定した変換テーブルに基づいて基本レイヤのLPC係数を拡張レイヤのLPC係数に変換し、拡張レイヤのLPC係数に基づいてスペクトル包絡を算出し、符号化処理におけるスペクトル正規化またはベクトル量子化の少なくとも一方に前記スペクトル包絡を活用する請求の範囲第5項に記載の符号化装置。
- 25 7. 前記拡張レイヤ符号化手段は、前記復号化手段の復号処理の中で生成されたピッチ周期およびピッチゲインを利用して符号化処理を行うことを特徴としたピッチ周期およびピッチゲインを利用して符号化処理を行うことを特徴とする

る請求の範囲第1項に記載の符号化装置。

8. 前記拡張レイヤ符号化手段は、ピッチ周期とピッチゲインを用いてスペクトル微細構造を算出し、符号化処理におけるスペクトル正規化およびベクトル量子化に前記スペクトル微細構造を活用する請求の範囲第7項に記載の符号化

5 装置。

9. 前記拡張レイヤ符号化手段は、前記復号化手段で生成された復号信号のパワーを利用して符号化処理を行うことを特徴とする請求の範囲第1項に記載の符号化装置。

10. 前記拡張レイヤ符号化手段は、復号信号のパワーに基づいてMDCT変換係数のパワーの変動量を量子化し、符号化処理におけるパワー正規化に前記15 量子化したMDCT変換係数のパワーの変動量を活用する請求の範囲第9項に記載の符号化装置。

11. 入力時の入力信号とサンプリングレートが上げられた復号信号との差分から誤差信号を得る減算手段と、サンプリングレートが上げられた復号信号に基づいて前記誤差信号を符号化する対象の周波数を決定する周波数決定手段と、15 を具備し、前記拡張レイヤ符号化手段は、前記周波数にある前記差分信号を符号化する請求項1に記載の音響符号化装置。

12. 聴覚に寄与しない振幅値を表す聴覚マスキングを算出する聴覚マスキング手段を具備し、前記拡張レイヤ符号化手段は、前記周波数決定手段において20 前記聴覚マスキング内の信号を符号化の対象としないように符号化する対象を決定して前記誤差信号のスペクトルである誤差スペクトルを符号化する請求の範囲第11項に記載の音響符号化装置。

13. 前記聴覚マスキング手段は、サンプリングレートが上げられた復号信号を周波数領域の係数に変換する周波数領域変換手段と、前記周波数領域の係数25 を用いて推定聴覚マスキングを算出する推定聴覚マスキング算出手段と、前記復号信号のスペクトルの振幅値が前記推定聴覚マスキングの振幅値を超える周

波数を求める決定手段と、を具備し、前記拡張レイヤ符号化手段は、前記周波数に位置する前記誤差スペクトルを符号化する請求の範囲第12項に記載の音響符号化装置。

14. 前記聴覚マスキング手段は、前記周波数領域の係数を用いて推定誤差ス

5 ペクトルを算出する推定誤差スペクトル算出手段を具備し、前記決定手段は、前記推定誤差スペクトルの振幅値が前記推定聴覚マスキングの振幅値を超える周波数を求める請求の範囲第13項に記載の音響符号化装置。

15. 前記聴覚マスキング手段は、前記推定聴覚マスキング算出手段において

10 算出された推定聴覚マスキングを平滑化する修正手段を具備し、前記決定手段は、前記復号信号のスペクトルまたは前記推定誤差スペクトルの振幅値が平滑化された前記推定聴覚マスキングの振幅値を超える周波数を求める特徴とする請求の範囲第13項に記載の音響符号化装置。

16. 前記拡張レイヤ符号化手段は、推定誤差スペクトルまたは誤差スペクト

ルのいずれかと聴覚マスキングまたは推定聴覚マスキングのいずれかとの振幅

15 値の差を周波数毎に算出し、前記振幅値の差の大きさに基づいて符号化の情報量を決定する請求の範囲第13項に記載の音響符号化装置。

17. 前記拡張レイヤ符号化手段は、前記決定手段で求めた周波数に加えて、

予め定められた帯域にある前記誤差スペクトルを符号化する請求の範囲第13

項に記載の音響符号化装置。

20 18. 符号化側において入力信号を所定の基本フレーム単位で符号化した第1

符号化コードを復号化して第1復号信号を得る基本レイヤ復号化手段と、第2

符号化コードを復号化して第2復号信号を得る拡張レイヤ復号化手段と、前記第1復号信号のサンプリングレートを前記第2復号信号と同一のレートに上げるアップサンプリング手段と、前記サンプリングレートを上げた第1復号信号

25 と前記第2復号信号とを加算する加算手段とを具備する復号化装置。

19. 前記基本レイヤ復号化手段は、符号励振線形予測法を用いて第1符号化

コードを復号化する請求の範囲第18項に記載の復号化装置。

20. 前記拡張レイヤ復号化手段は、直交変換を用いて第2符号化コードを復号化する請求の範囲第18項に記載の復号化装置。

21. 前記拡張レイヤ復号化手段は、逆MDCT変換を用いて第2符号化コードを復号化する請求の範囲第20項に記載の復号化装置。

22. 前記拡張レイヤ復号化手段は、基本レイヤのLPC係数を利用して第2符号化コードを復号化する請求の範囲第18項に記載の復号化装置。

23. 前記拡張レイヤ復号化手段は、予め設定した変換テーブルに基づいて基本レイヤのLPC係数を拡張レイヤのLPC係数に変換し、拡張レイヤのLPC係数に基づいてスペクトル包絡を算出し、復号化処理におけるベクトル復号化に前記スペクトル包絡を活用する請求の範囲第22項に記載の復号化装置。

24. 前記拡張レイヤ復号化手段は、ピッチ周期またはピッチゲインの少なくとも一方を利用して復号化処理を行うことを特徴とする請求の範囲第18項に記載の復号化装置。

25. 前記拡張レイヤ復号化手段は、ピッチ周期とピッチゲインを用いてスペクトル微細構造を算出し、復号化処理におけるベクトル復号化に前記スペクトル微細構造を活用する請求の範囲第24項に記載の復号化装置。

26. 前記拡張レイヤ復号化手段は、前記復号化手段で生成された復号信号のパワーを利用して復号化処理を行うことを特徴とする請求の範囲第18項に記載の復号化装置。

27. 前記拡張レイヤ復号化手段は、復号信号のパワーに基づいてMDCT変換係数のパワーの変動量を復号化し、復号化処理におけるパワー正規化に前記復号化したMDCT変換係数のパワーの変動量を活用する請求の範囲第26項に記載の復号化装置。

28. 前記アップサンプリングされた第1復号信号に基づいて入力信号と符号化側において第1符号化コードを復号した信号との残差信号を符号化した第2

符号化コードを復号化する対象の周波数を決定する周波数決定手段を具備し、前記拡張レイヤ復号化手段は、前記周波数の情報を用いて前記第2符号化コードを復号化して第2復号信号を生成し、前記加算手段は、前記第2復号信号とサンプリングレートを上げられた第1復号信号を加算する請求の範囲第18項

5 に記載の音響復号化装置。

29. 聴覚に寄与しない振幅値を表す聴覚マスキングを算出する聴覚マスキング手段を具備し、前記拡張レイヤ復号化手段は、前記周波数決定手段において前記聴覚マスキング内の信号を復号化の対象としないように復号化する対象を決定する請求の範囲第28項に記載の音響復号化装置。

10 30. 前記聴覚マスキング手段は、サンプリングレートの上げられた基本レイヤの復号信号を周波数領域の係数に変換する周波数領域変換手段と、前記周波数領域の係数を用いて推定聴覚マスキングを算出する推定聴覚マスキング算出手段と、前記復号信号のスペクトルの振幅値が前記推定聴覚マスキングの振幅値を超える周波数を求める決定手段と、を具備し、前記拡張レイヤ復号化手段は、前記周波数に位置する前記誤差スペクトルを復号化する請求の範囲第29項に記載の音響復号化装置。

31. 前記聴覚マスキング手段は、前記周波数領域の係数を用いて推定誤差スペクトルを算出する推定誤差スペクトル算出手段を具備し、前記決定手段は、前記推定誤差スペクトルの振幅値が前記推定聴覚マスキングの振幅値を超える周波数を求ることを特徴とする請求の範囲第30項に記載の音響復号化装置。

32. 前記聴覚マスキング手段は、前記推定聴覚マスキング算出手段において算出された推定聴覚マスキングを平滑化する修正手段を具備し、前記決定手段は、前記復号信号のスペクトルまたは前記推定誤差スペクトルの振幅値が平滑化された前記推定聴覚マスキングの振幅値を超える周波数を求ることを特徴

25 とする請求の範囲第30項に記載の音響復号化装置。

33. 前記拡張レイヤ復号化手段は、推定誤差スペクトルまたは誤差スペクト

ルのいずれかと聴覚マスキングまたは推定聴覚マスキングのいずれかとの振幅値の差を周波数毎に算出し、前記振幅値の差の大きさに基づいて復号化の情報量を決定する請求の範囲第29項に記載の音響復号化装置。

34. 前記拡張レイヤ復号化手段は、前記決定手段で求めた周波数に加えて、  
5 予め定められた帯域にある前記誤差スペクトルと、を復号化する請求の範囲第  
29項に記載の音響復号化装置。

35. 音響信号を電気的信号に変換する音響入力手段と、この音響入力手段から出力された信号をデジタル信号に変換するA/D変換手段と、このA/D変換手段から出力されたデジタル信号を符号化する請求の範囲第1項に記載の符号化装置と、この符号化装置から出力された符号化コードを無線周波数の信号に変調するRF変調手段と、このRF変調手段から出力された信号を電波に変換して送信する送信アンテナとを具備する音響信号送信装置。

36. 電波を受信する受信アンテナと、この受信アンテナに受信された信号を復調するRF復調手段と、このRF復調手段にて得られた情報を復号する請求の範囲第18項に記載の復号化装置と、この復号化装置から出力された信号をアナログ信号に変換するD/A変換手段と、このD/A変換手段から出力された電気的信号を音響信号に変換する音響出力手段とを具備する音響信号受信装置。

37. 請求の範囲35記載の音響信号送信装置を具備する通信端末装置。  
20 38. 請求の範囲36記載の音響信号受信装置を具備する通信端末装置。  
39. 請求の範囲35記載の音響信号送信装置を具備する基地局装置。

40. 請求の範囲36記載の音響信号受信装置を具備する基地局装置。  
41. 入力信号のサンプリングレートを下げる工程と、サンプリングレートを下げた入力信号を符号化して第1符号化コードを得る工程と、前記第1符号化  
25 コードに基づいて復号信号を生成する工程と、前記復号信号のサンプリングレートを前記入力信号と同一のレートに上げる工程と、前記復号信号を生成する

処理の中で得られたパラメータを利用し、前記入力信号と前記サンプリングレートを上げた復号信号との差分値を符号化して第2符号化コードを得る工程と、前記第1符号化コードと前記第2符号化コードとを多重する工程とを具備する符号化方法。

- 5 4 2. 第1符号化コードを復号化して第1復号信号を得る工程と、第2符号化コードを復号化して第2復号信号を得る工程と、前記第1復号信号のサンプリングレートを前記第2復号信号と同一のレートに上げる工程と、前記サンプリングレートを上げた第1信号と前記第2信号とを加算する工程とを具備する復号化方法。

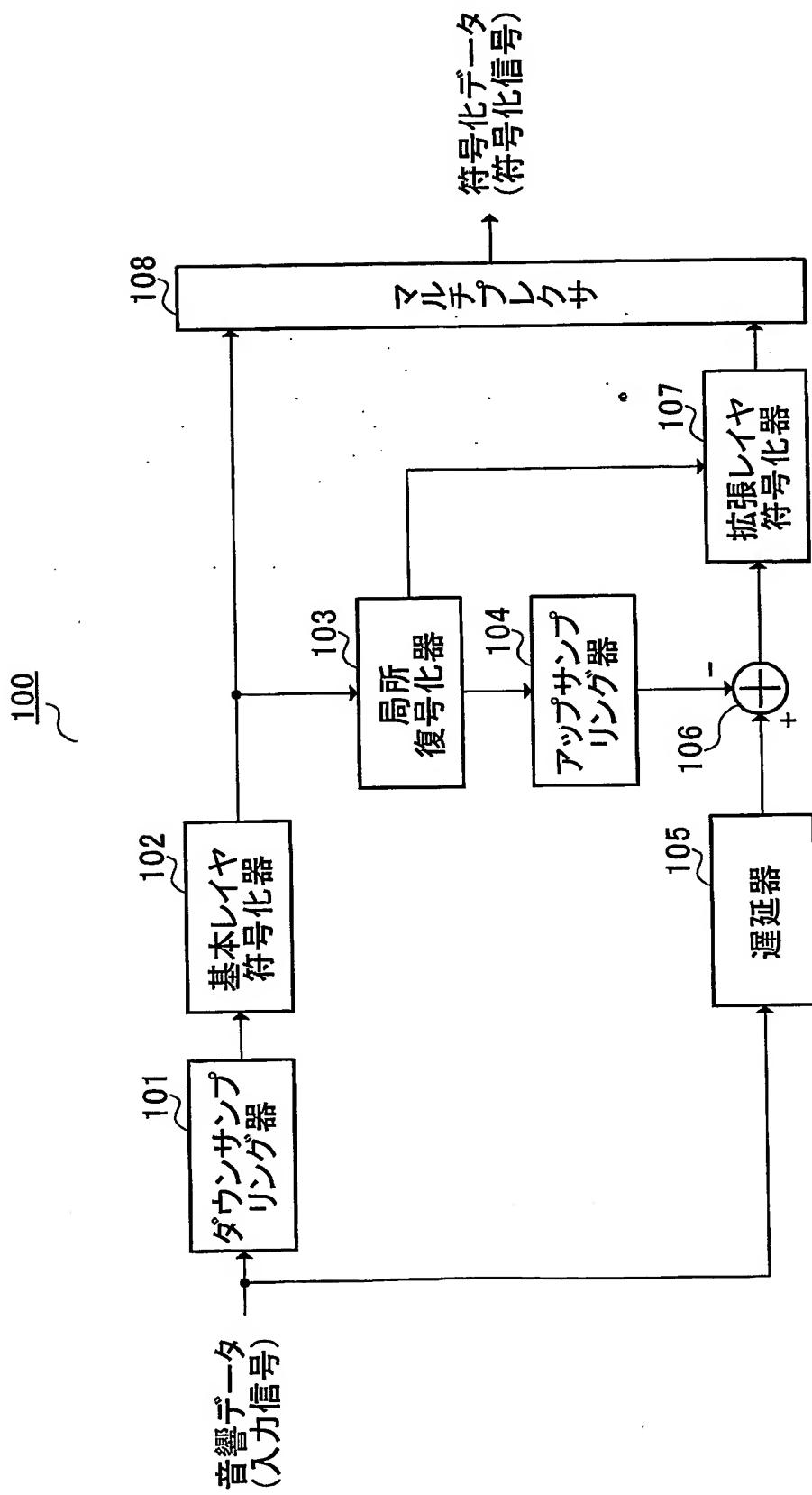


図 1

2 / 39

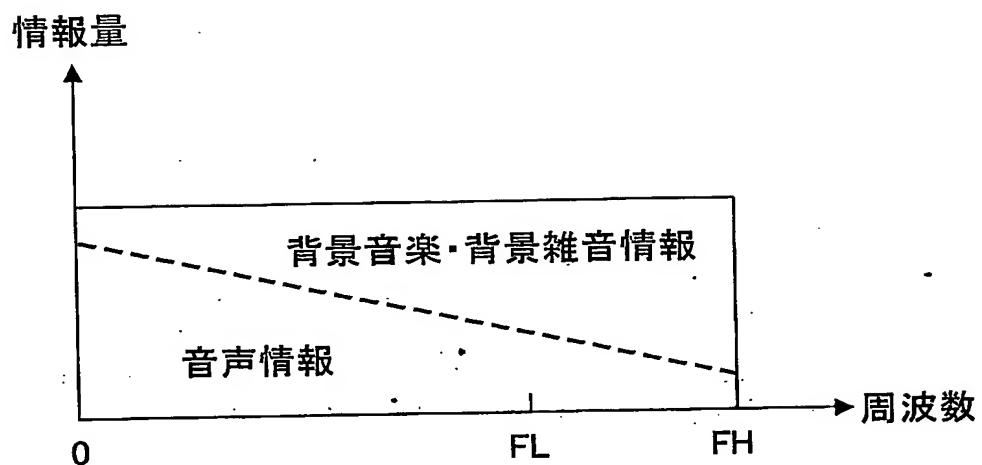


図 2

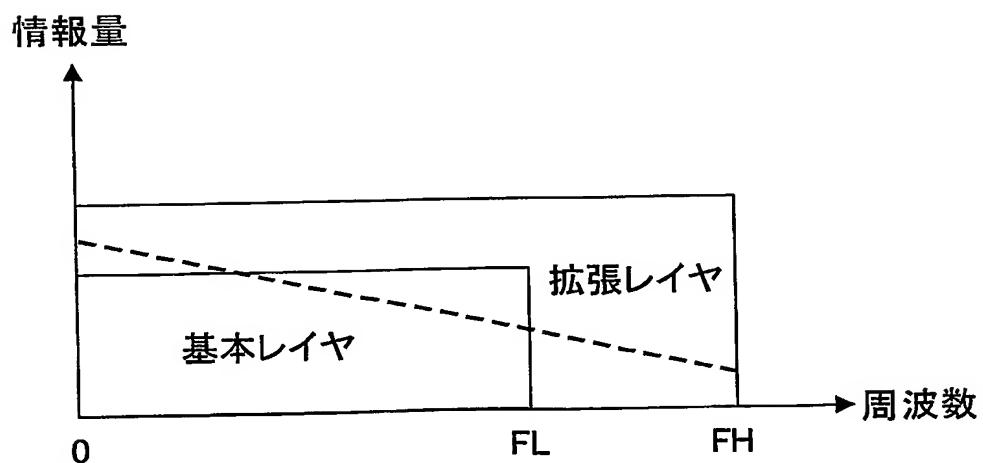


図 3

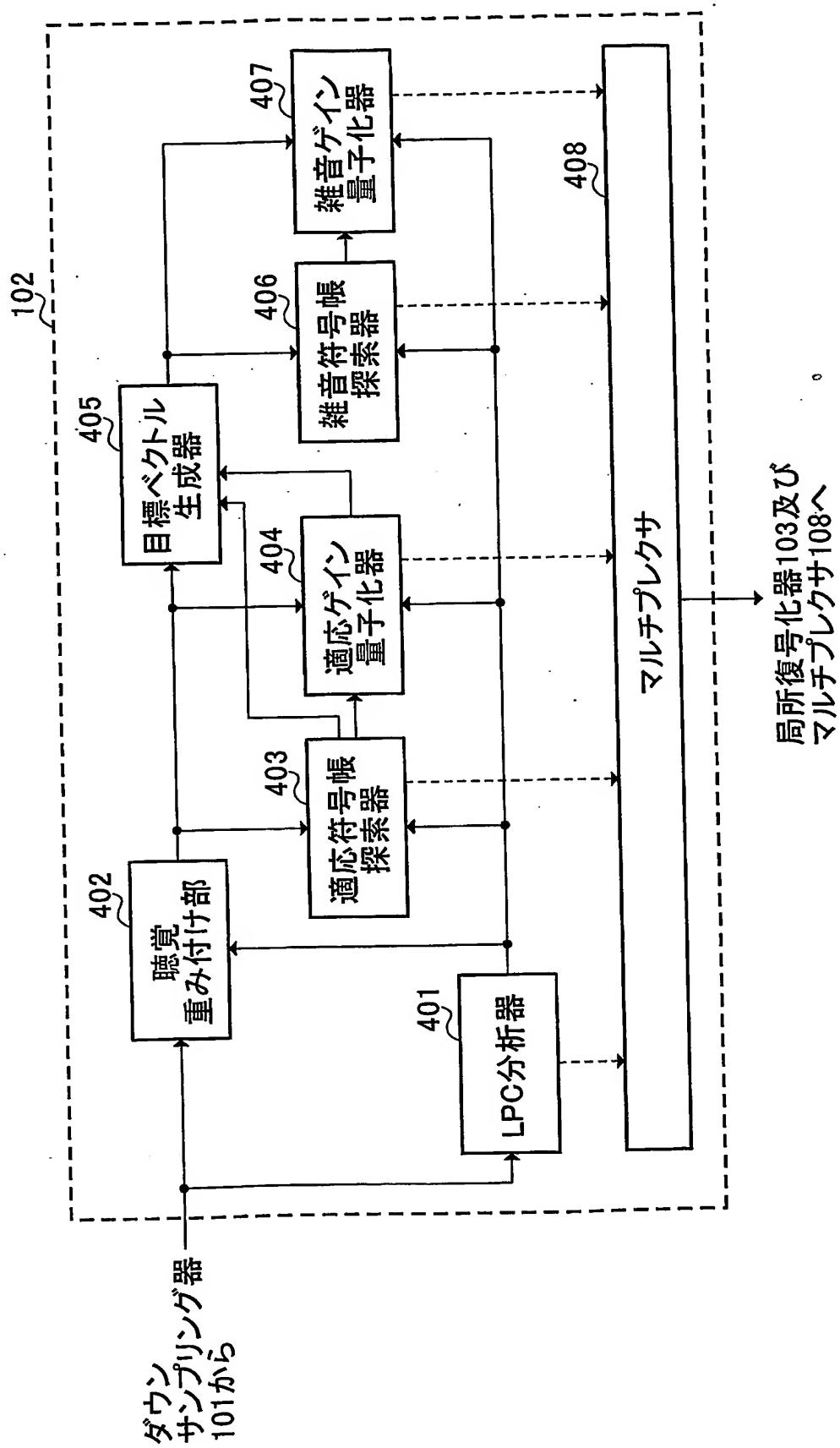


図 4

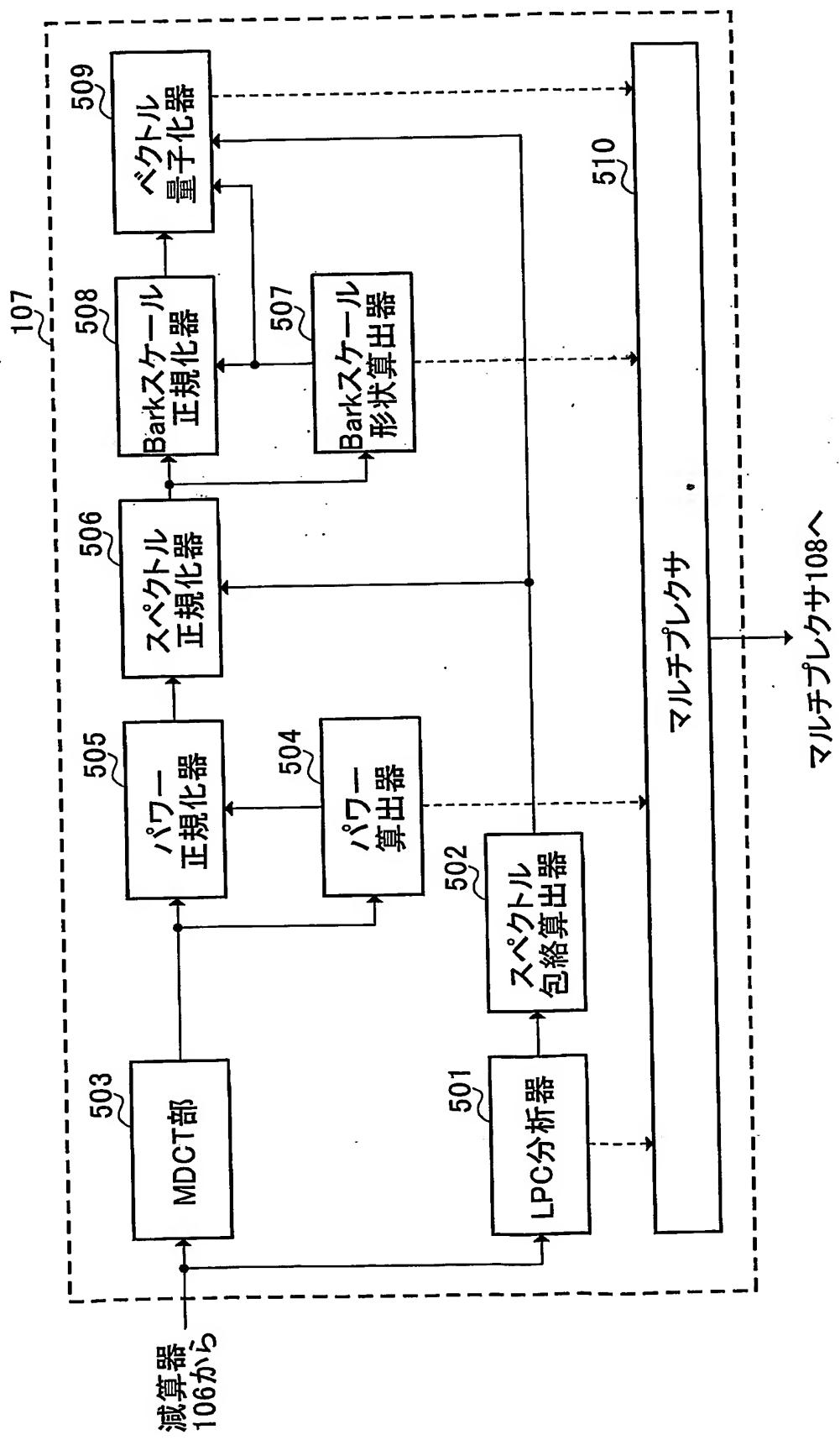


図 5

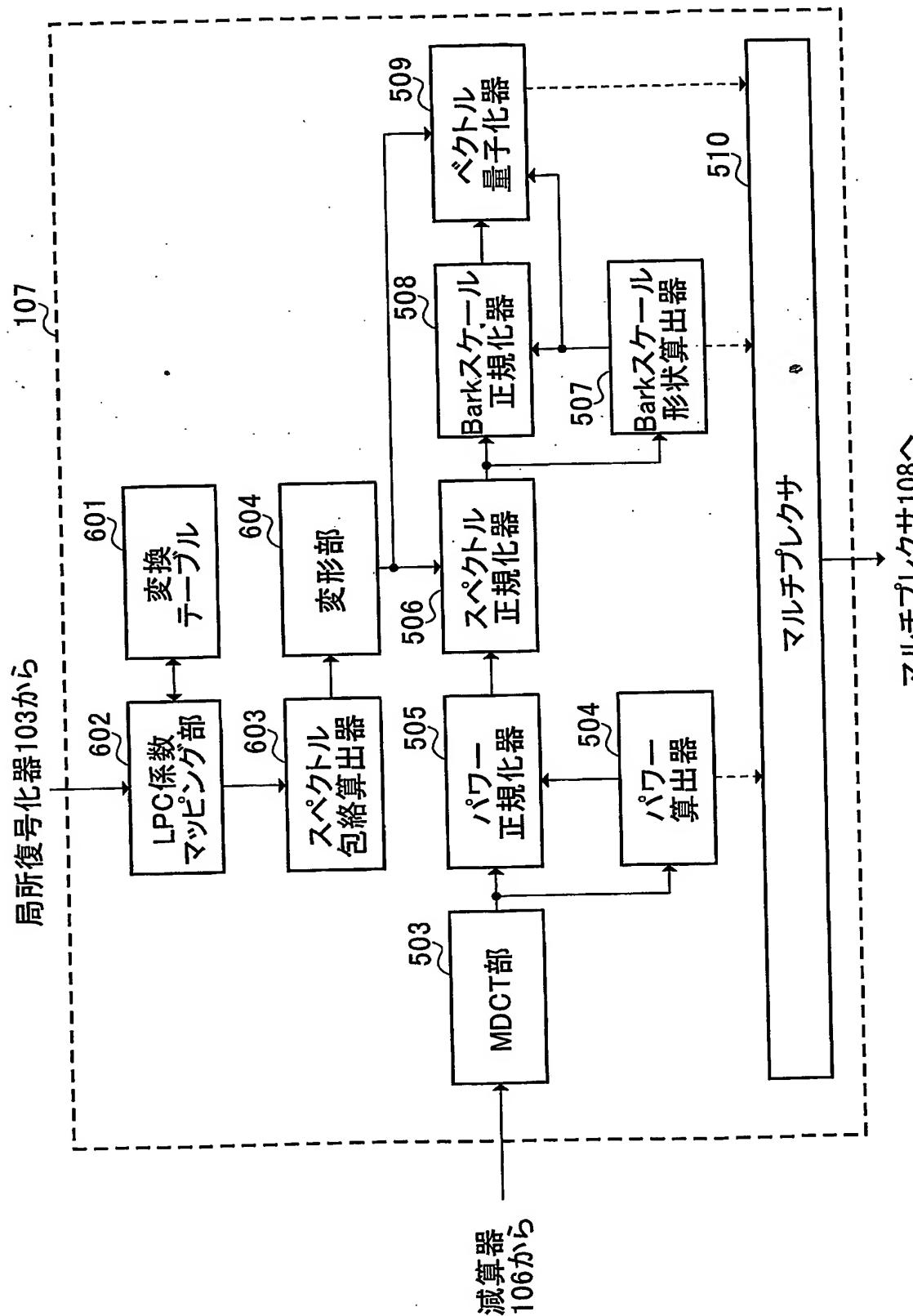


図 6

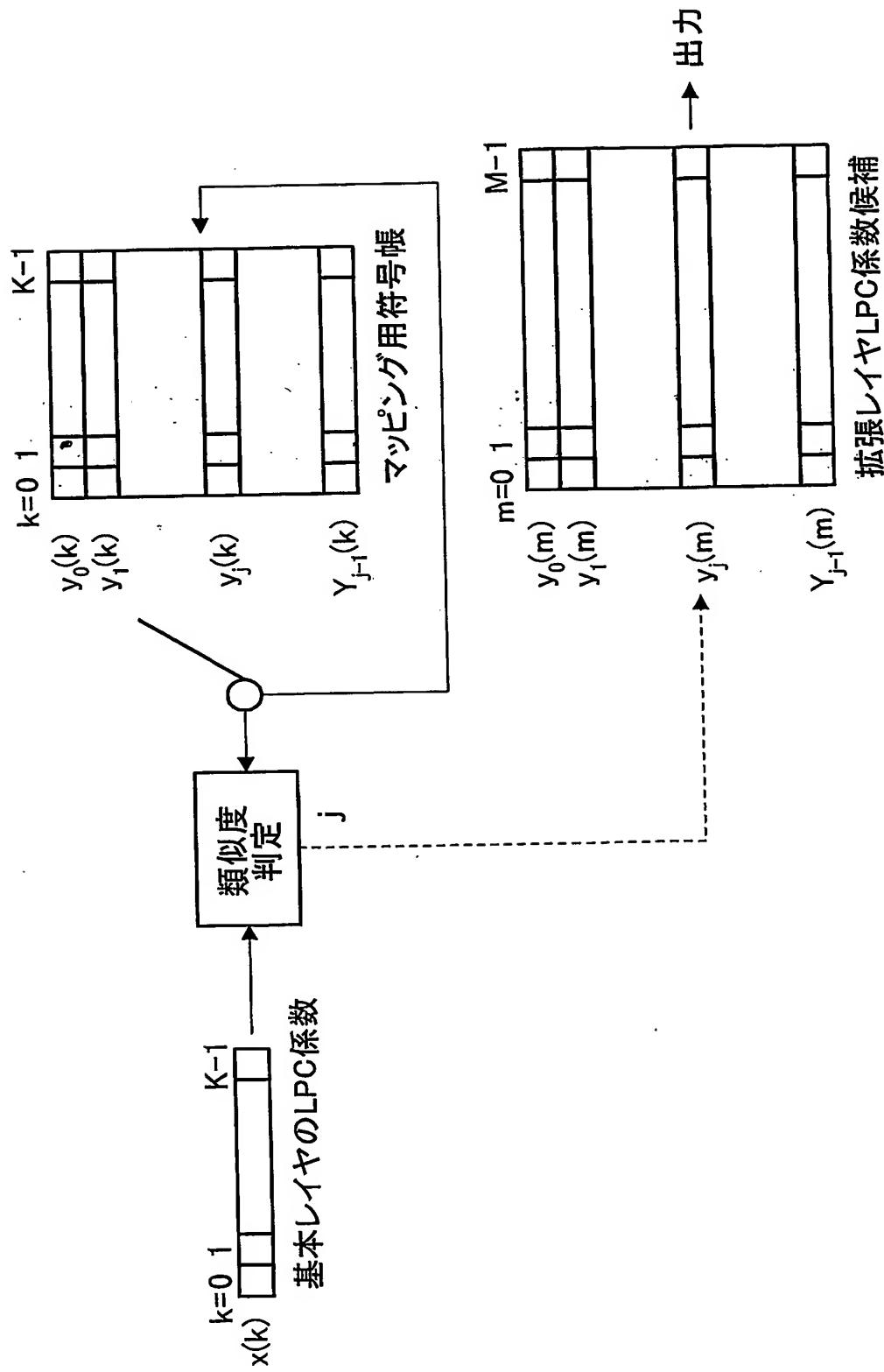


図7

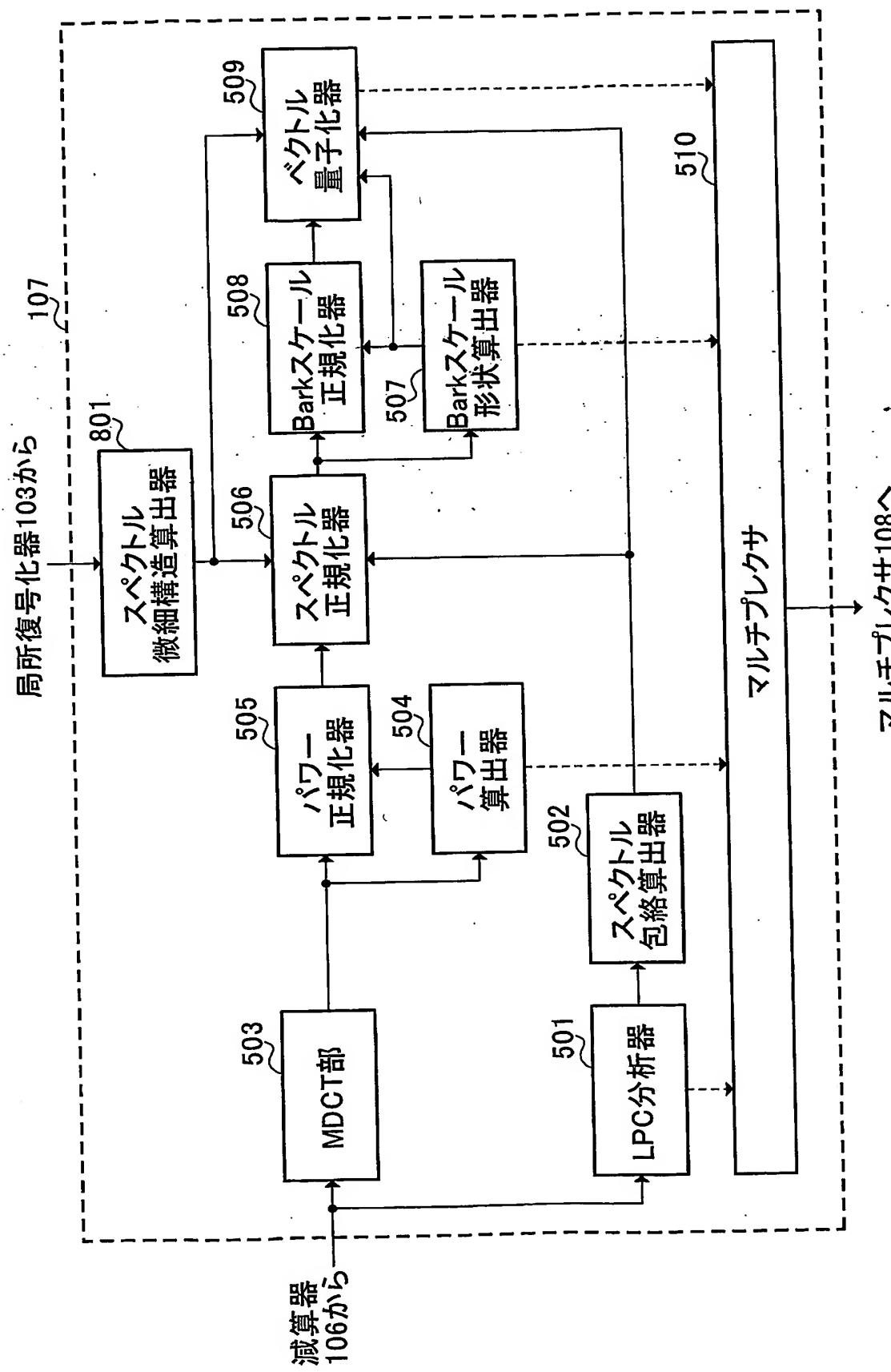


図 8

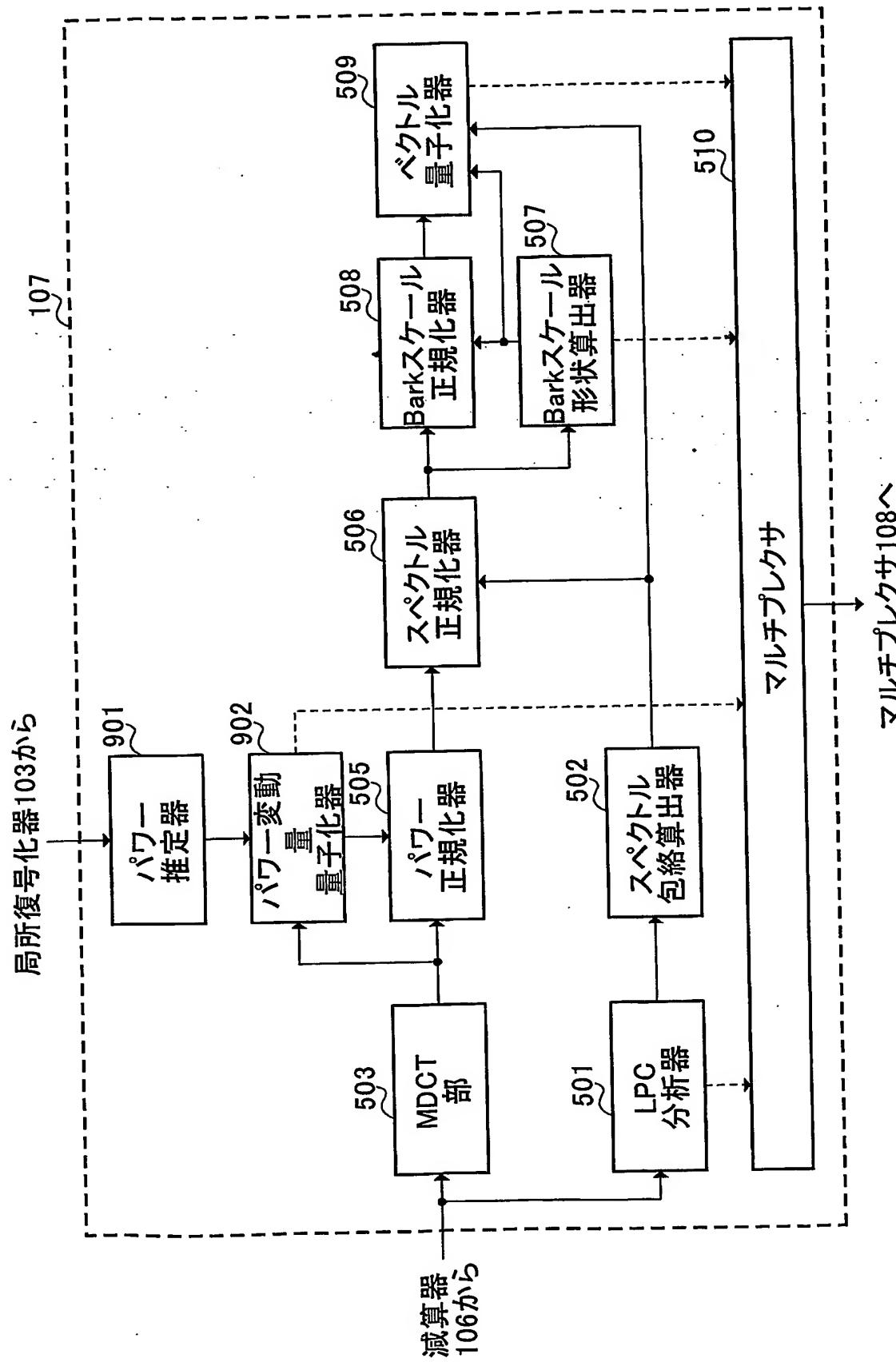


図 9

9 / 39

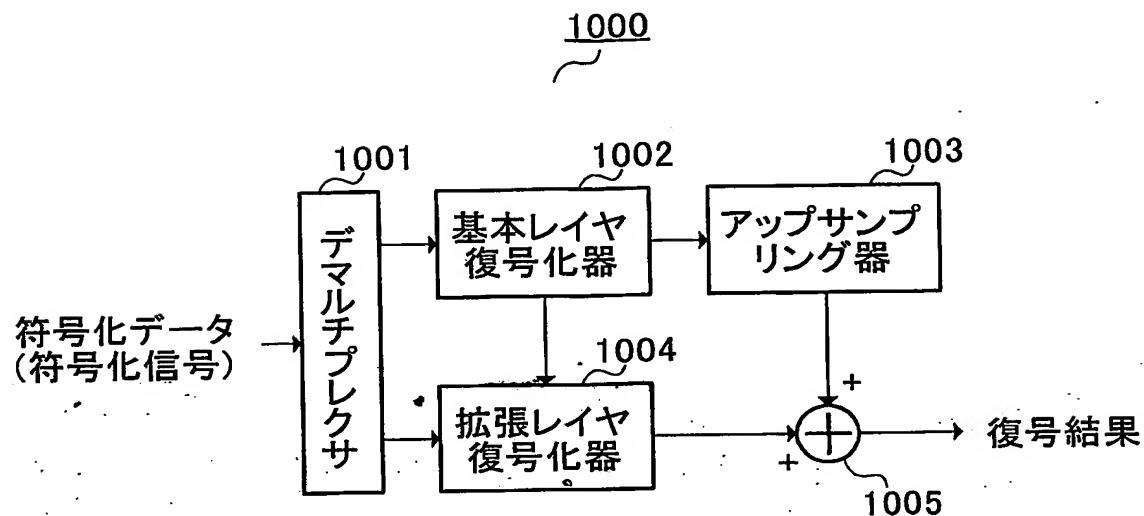


図 10

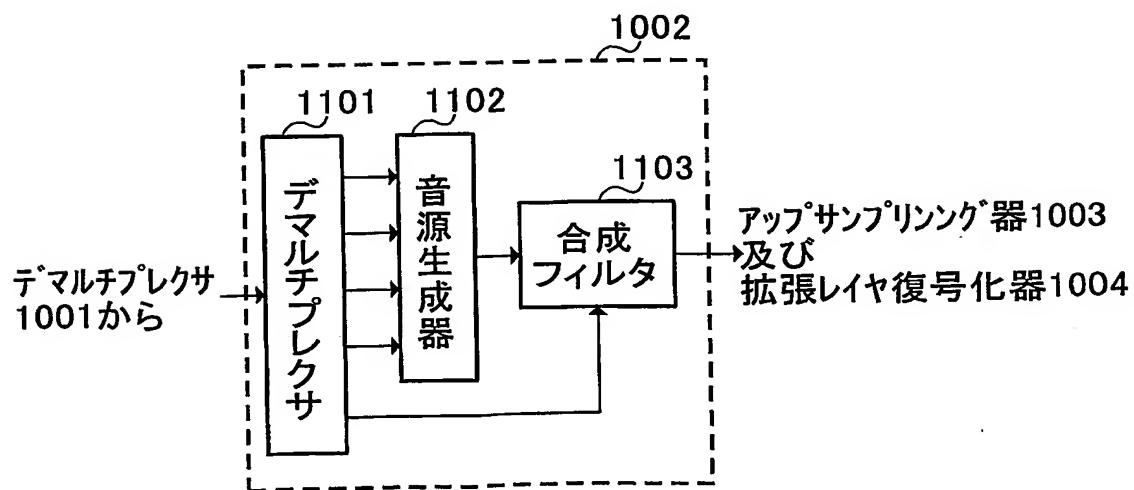


図 11

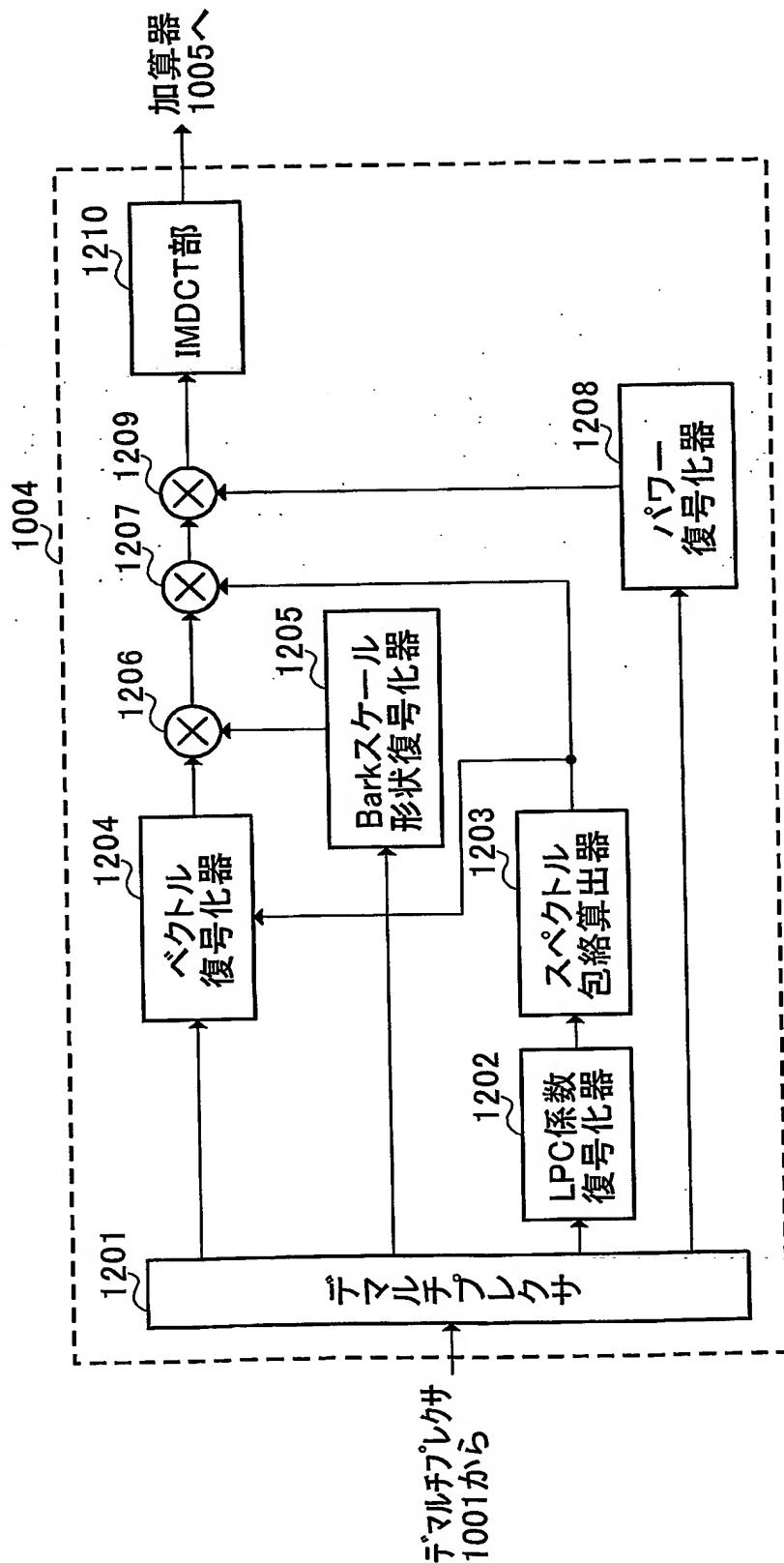


図 12

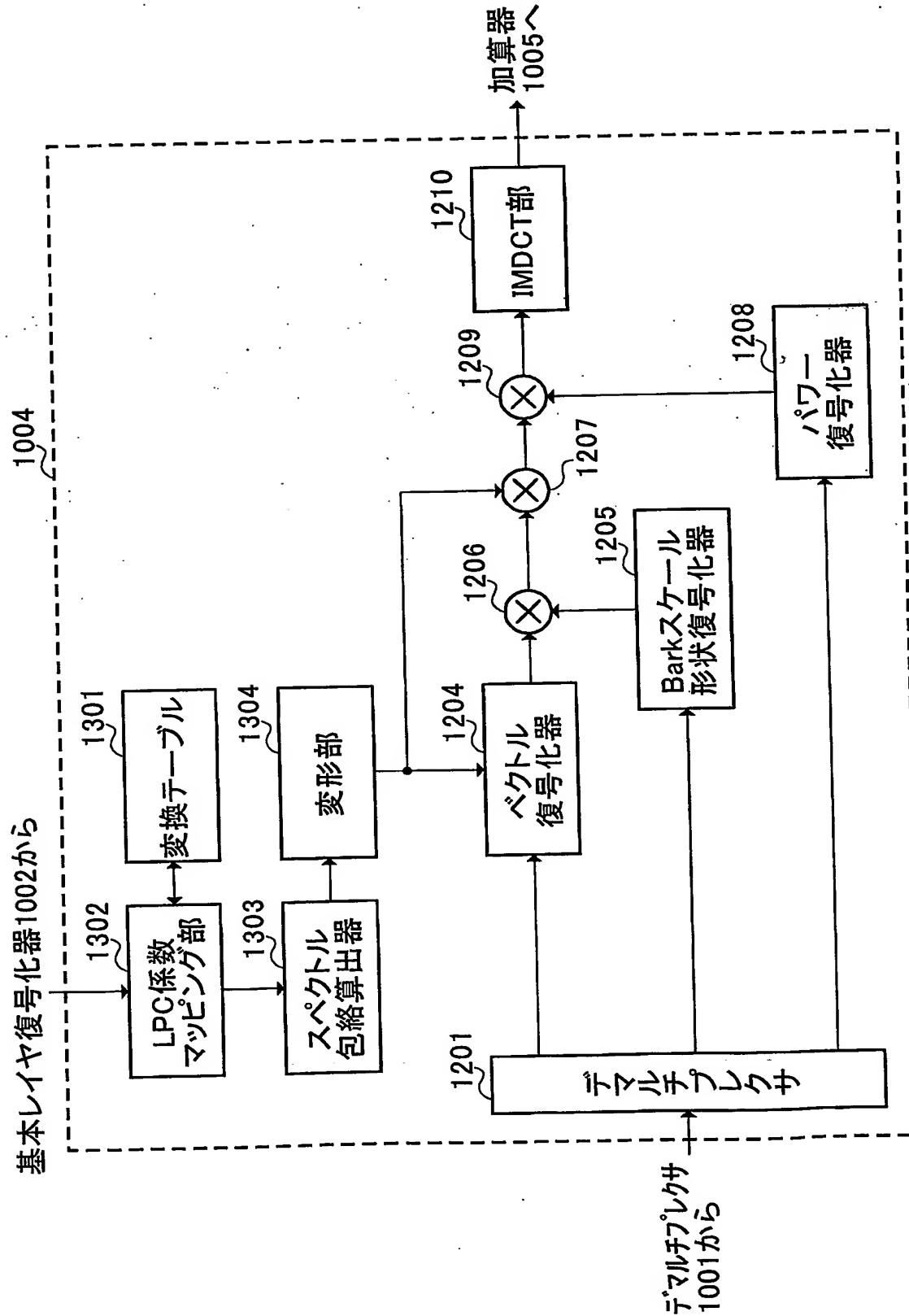


図13

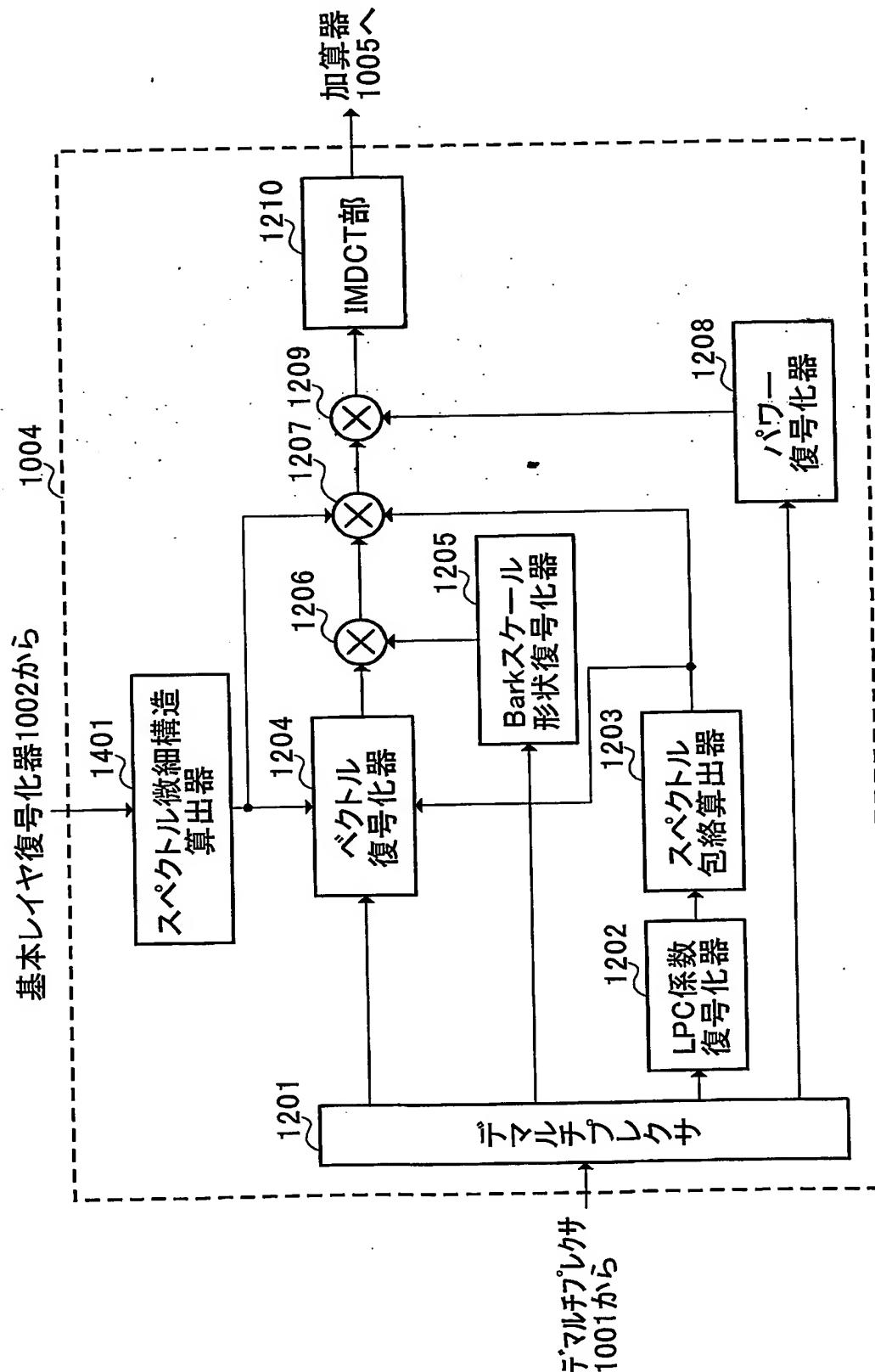


図 14

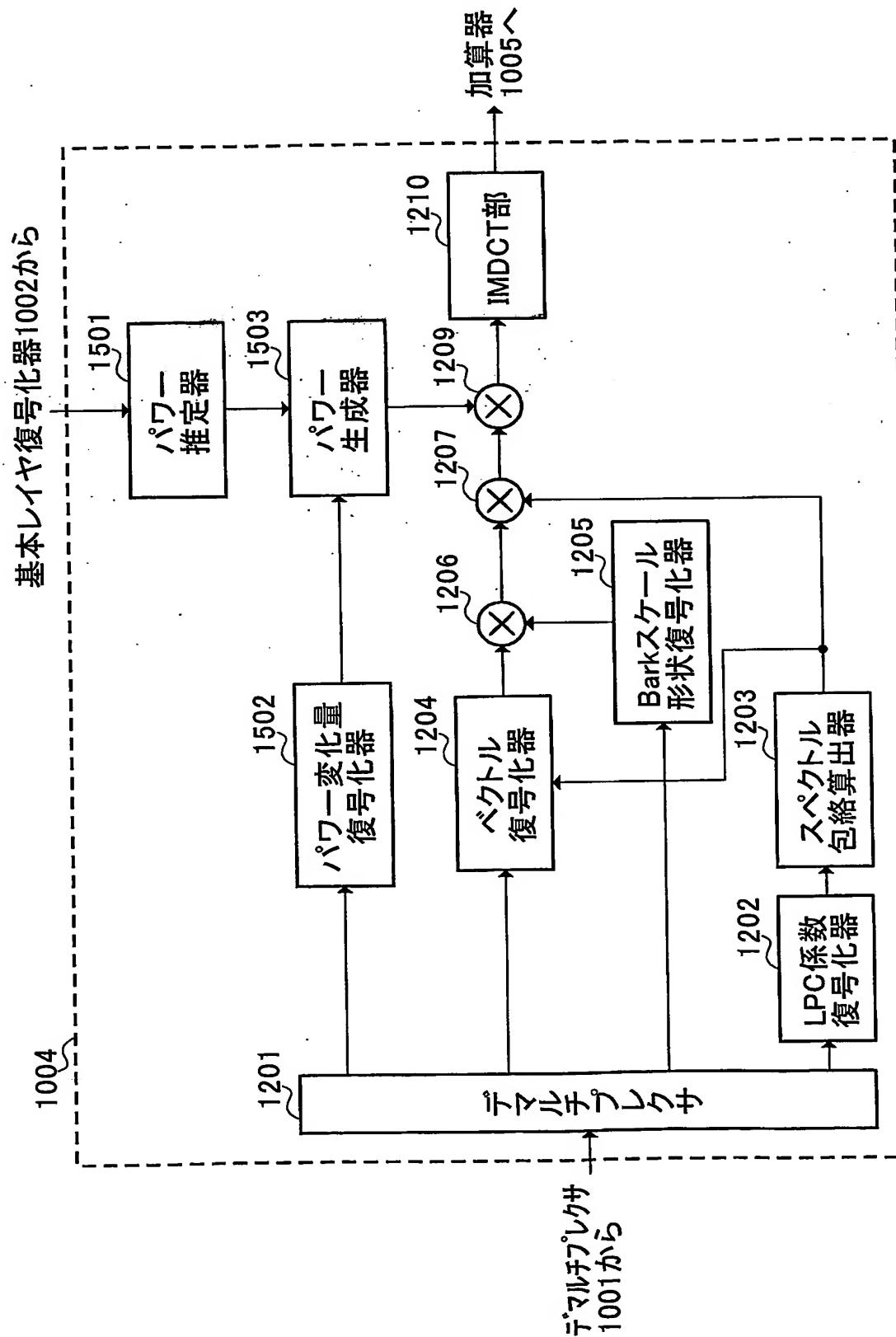


図 15

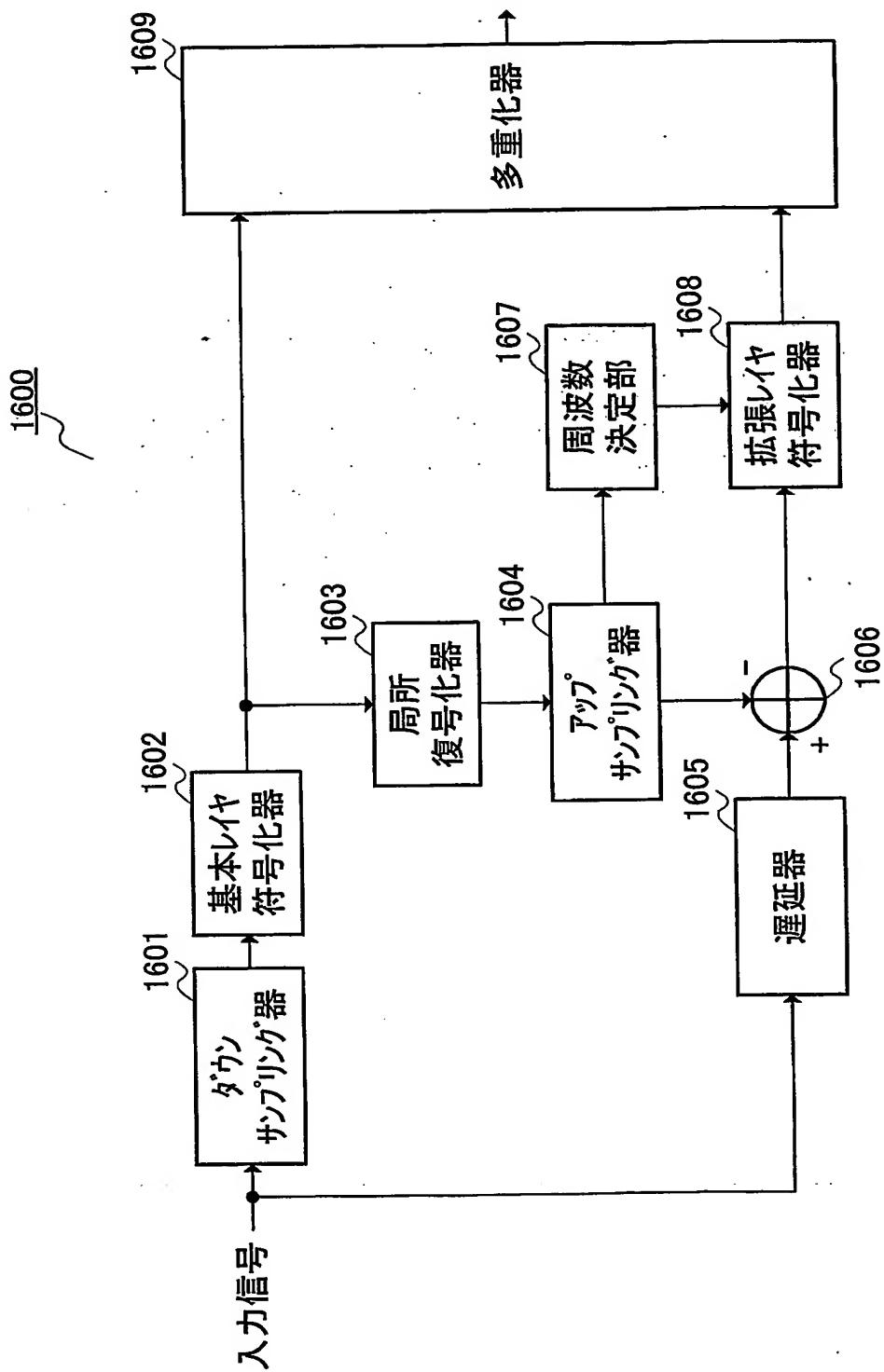


図16

15/39

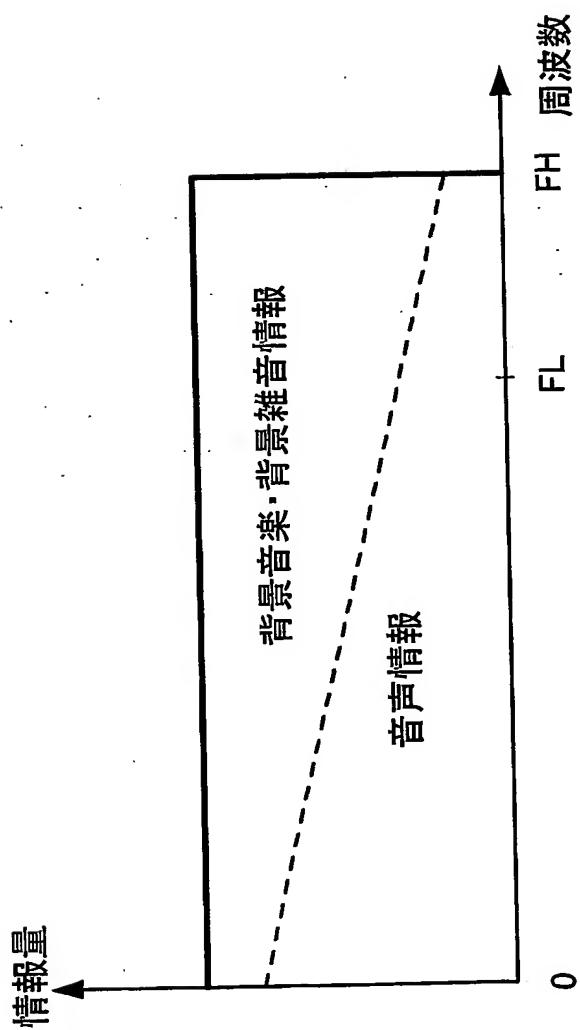


図17

16/39

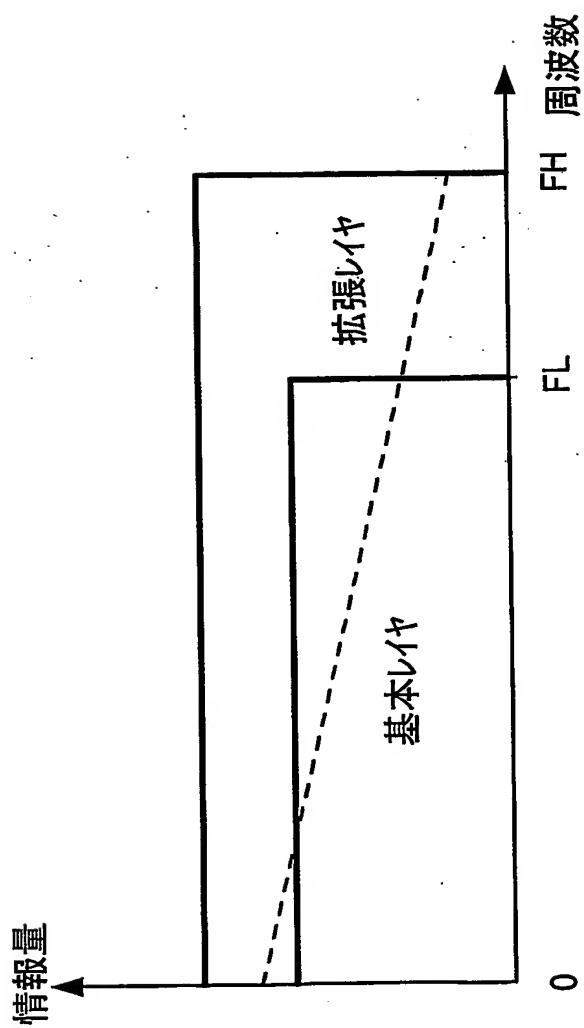


図18

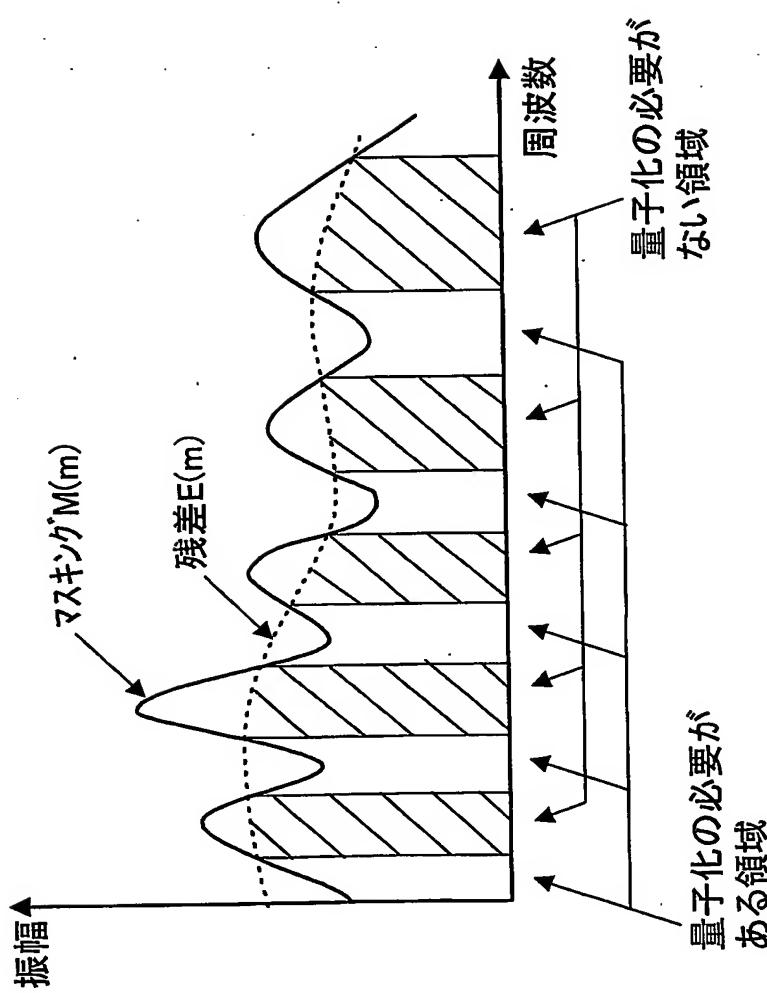


図19

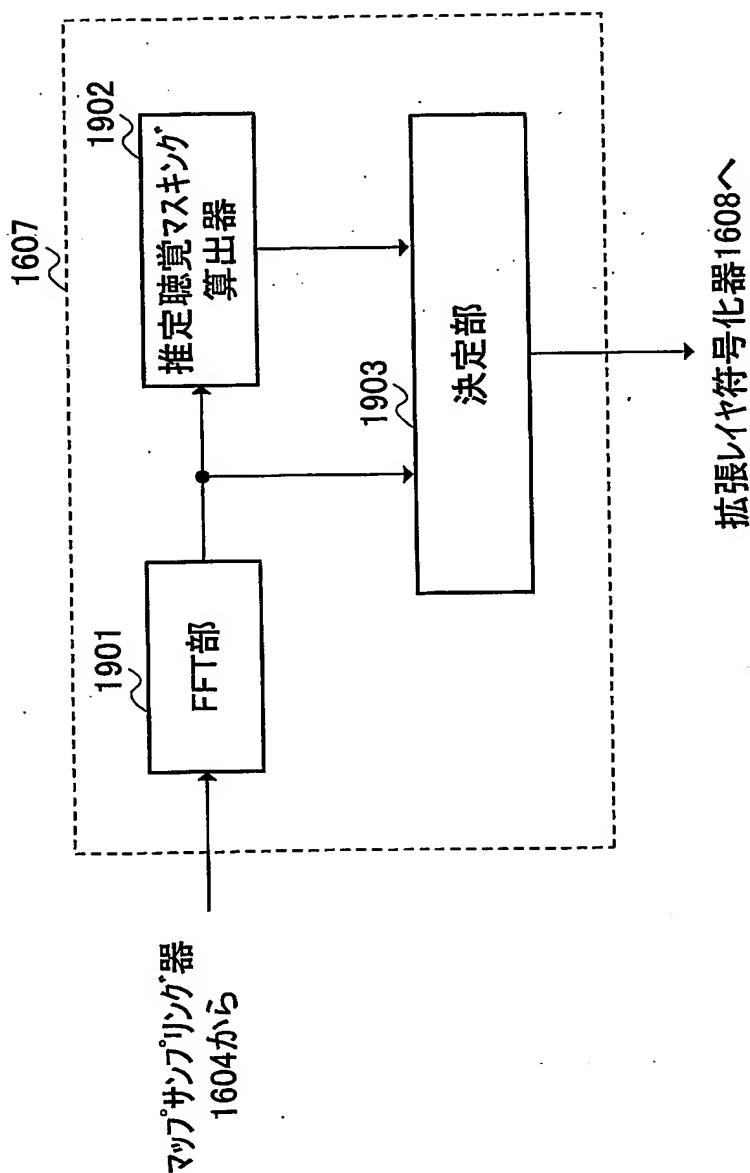


図20

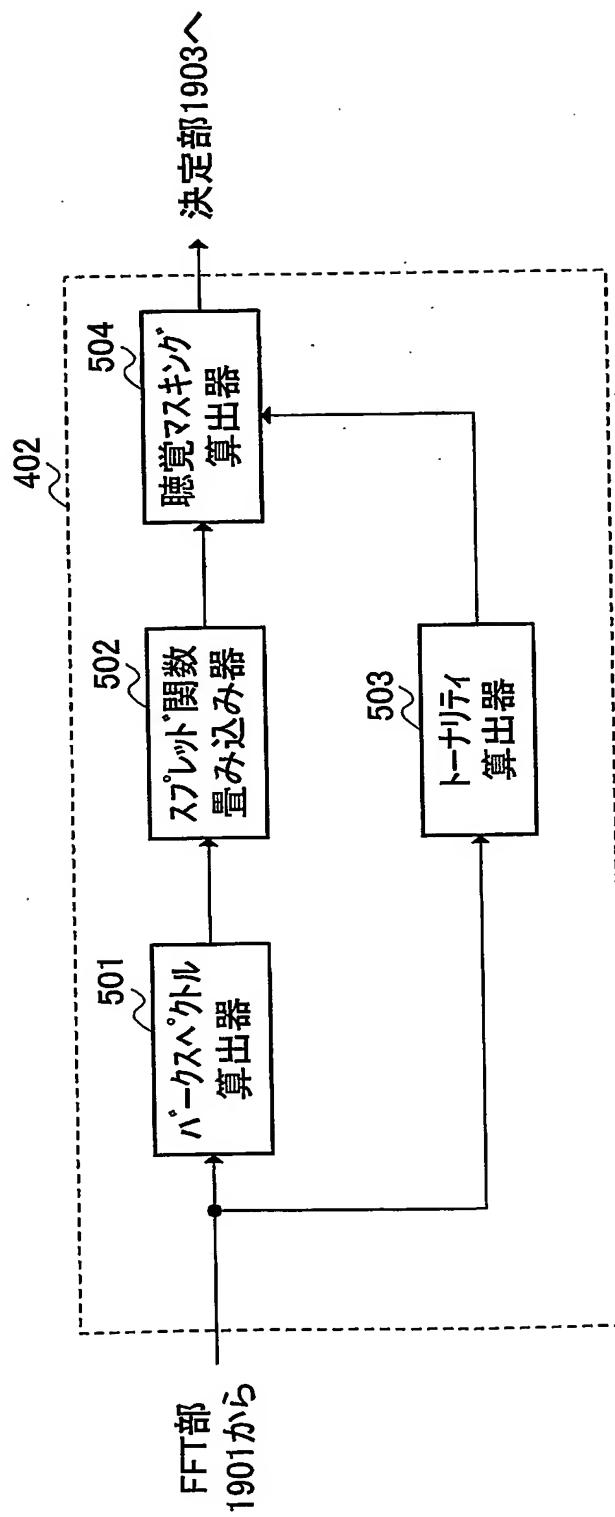


図21

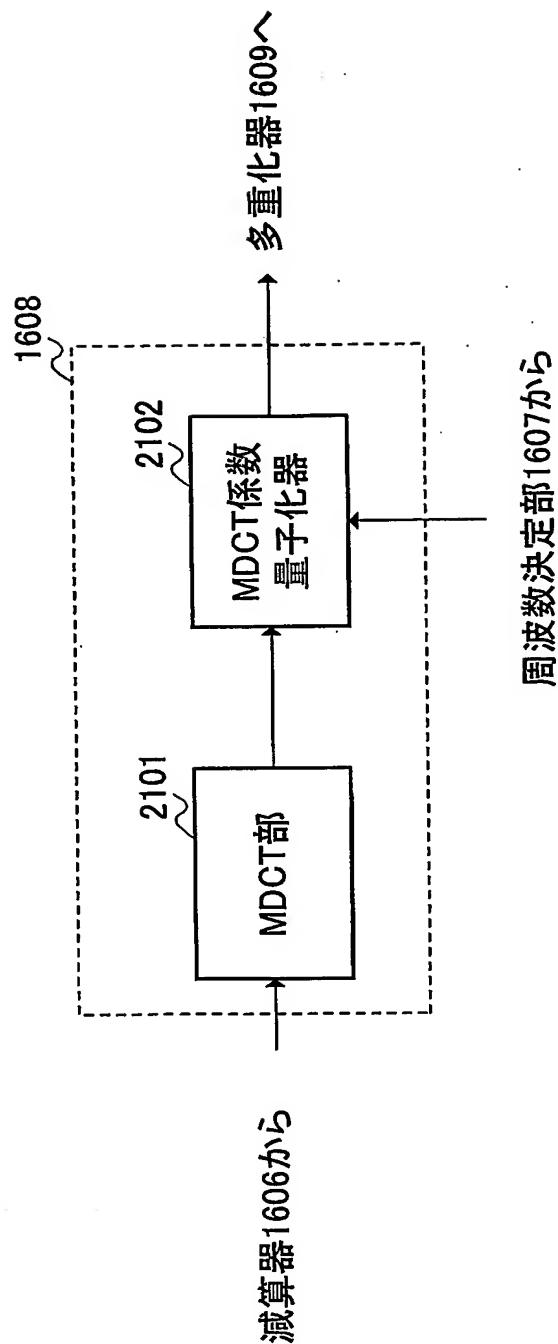


図22

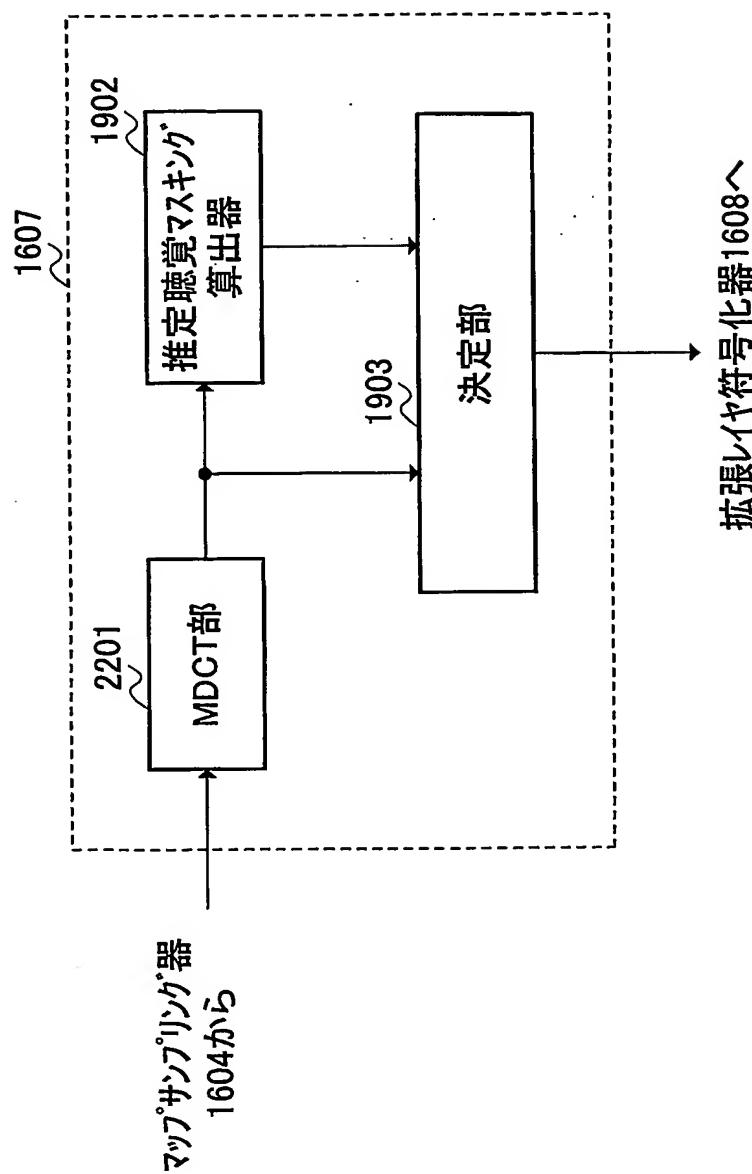


図23

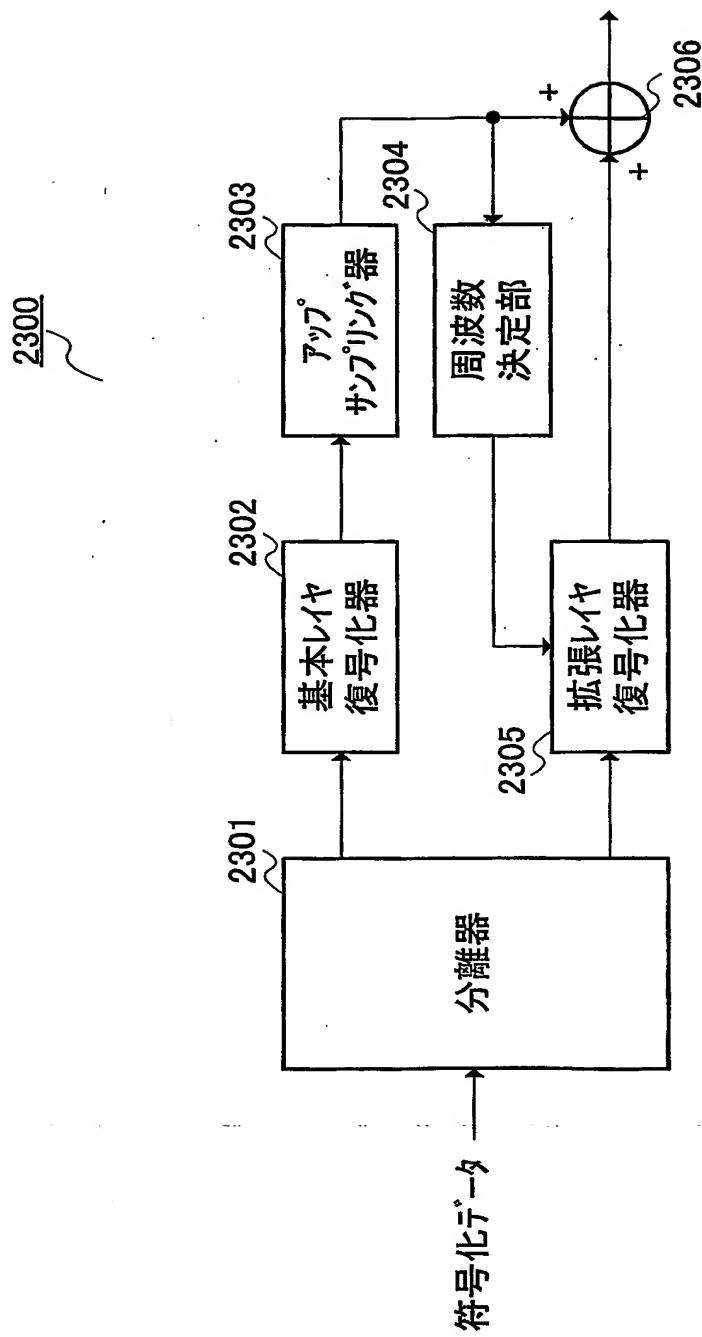
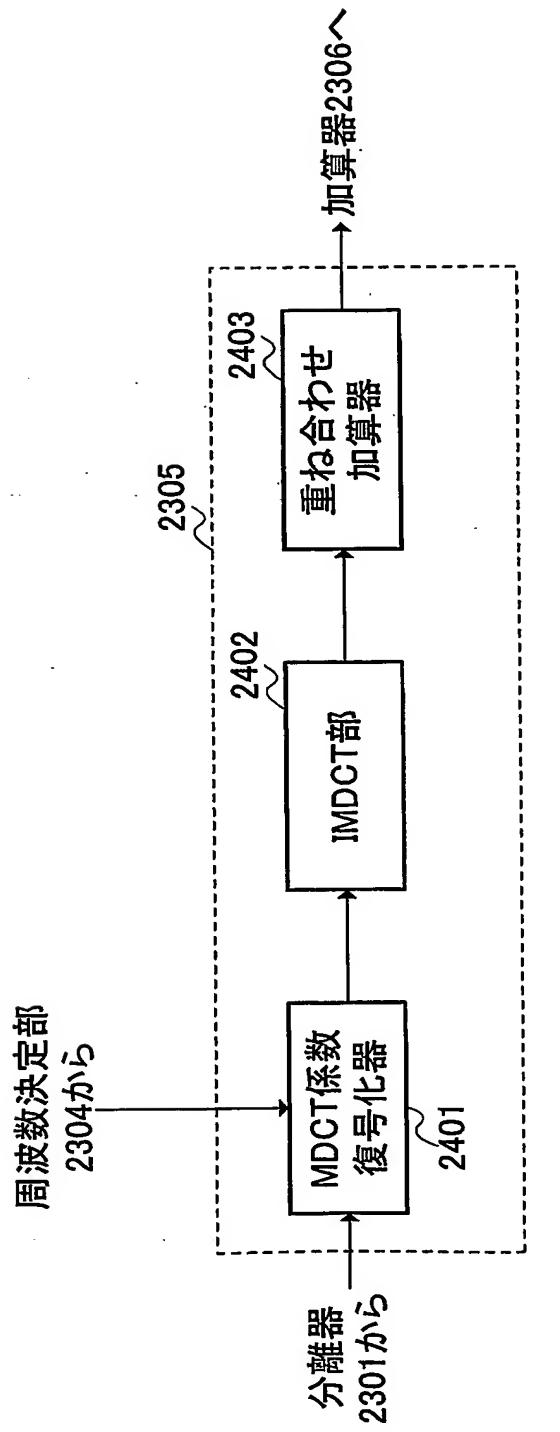


図24



25

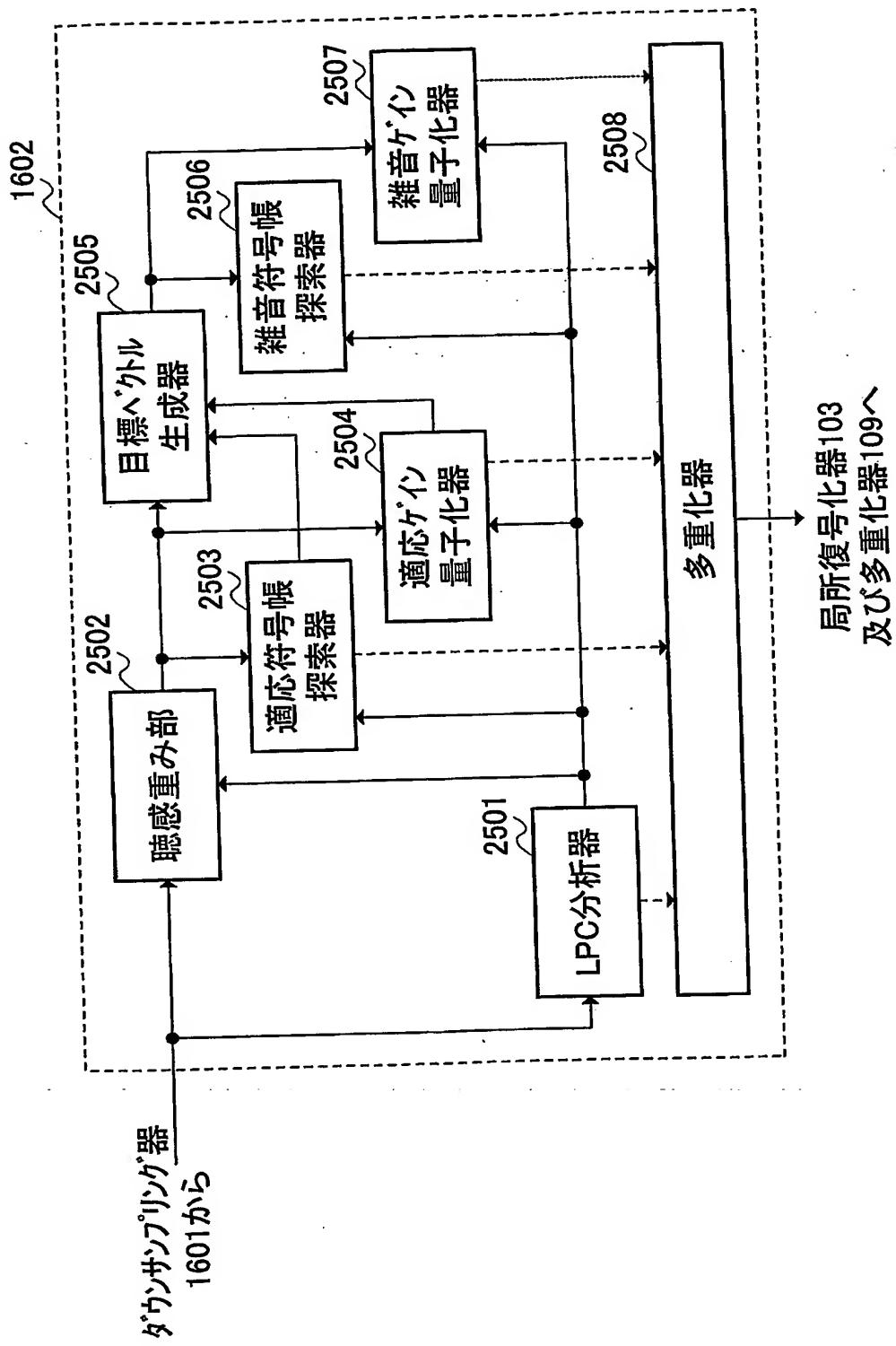


図26

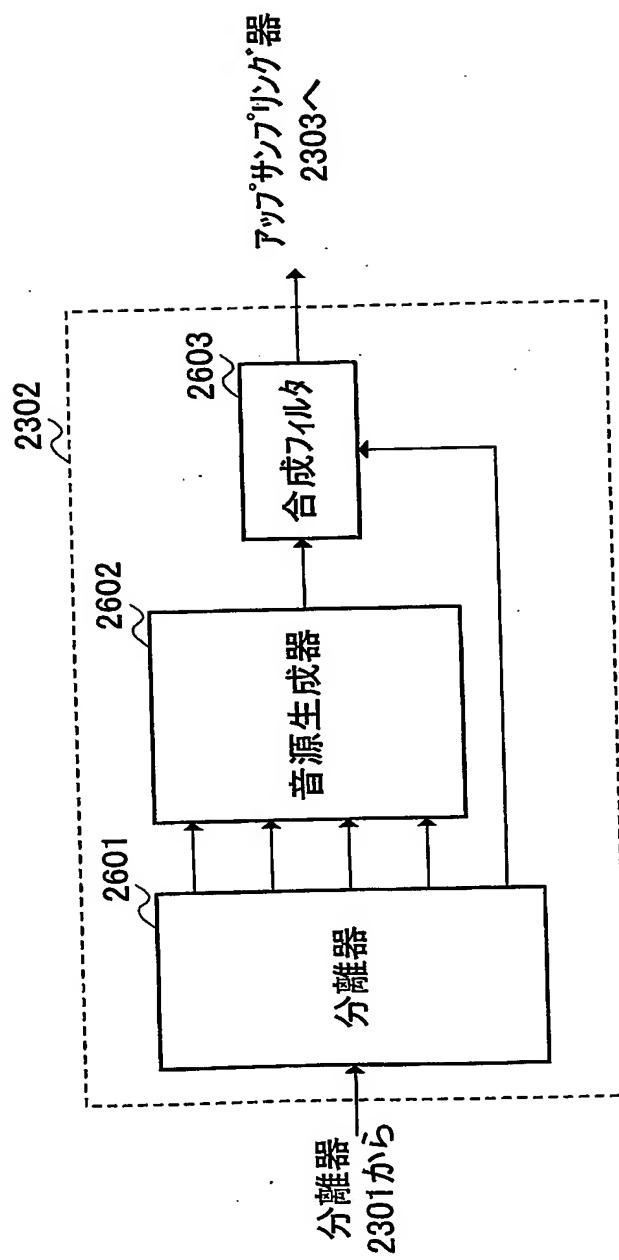


図27

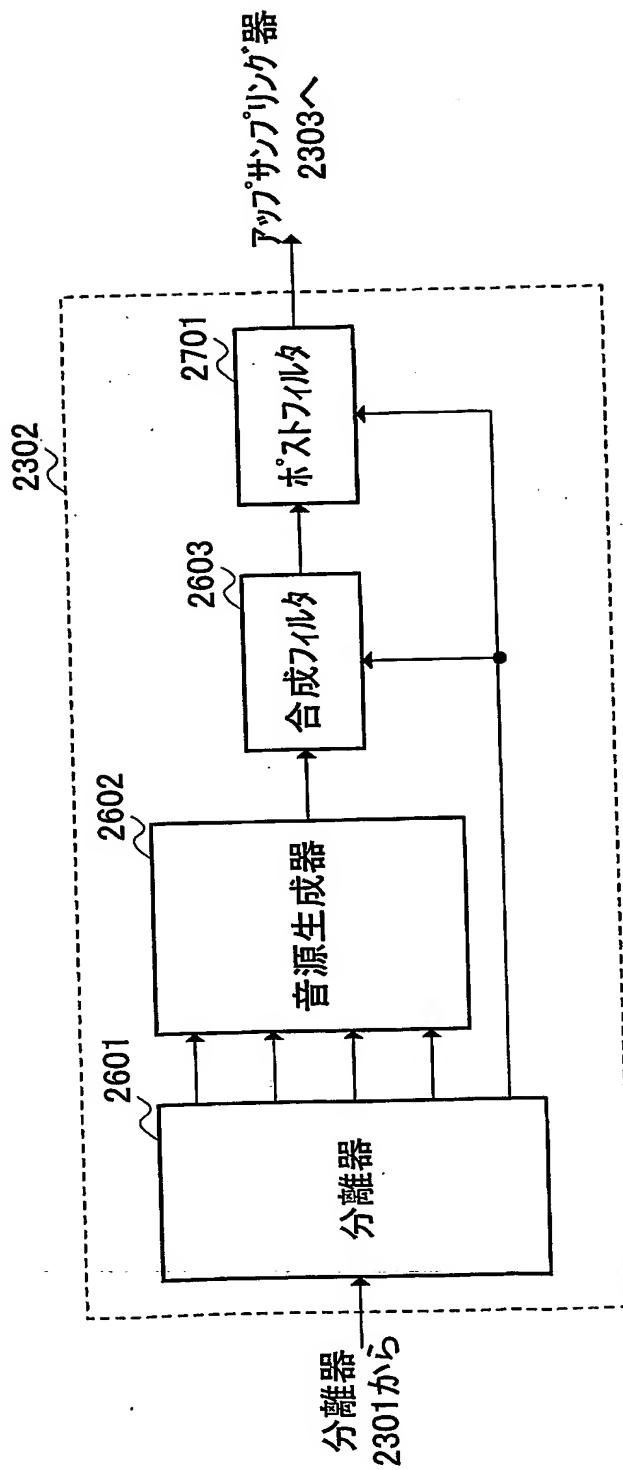


図28

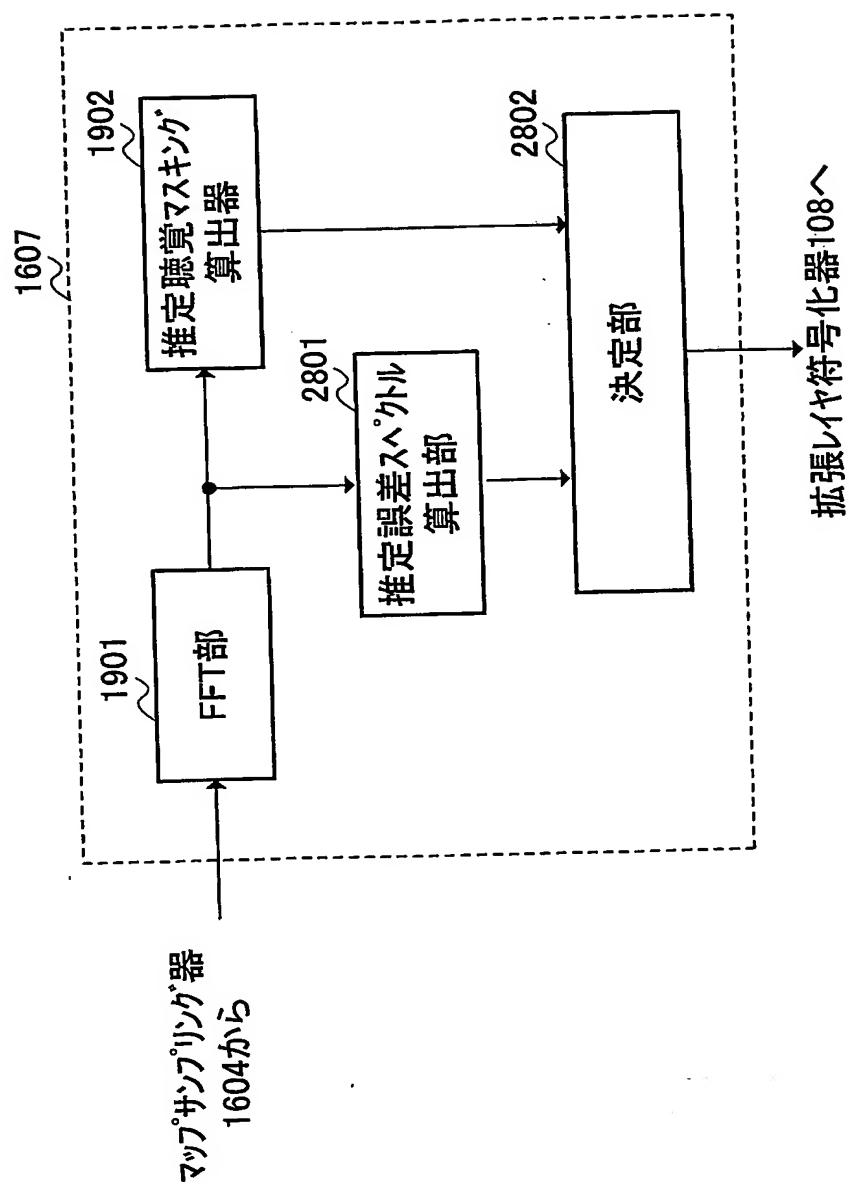


図29

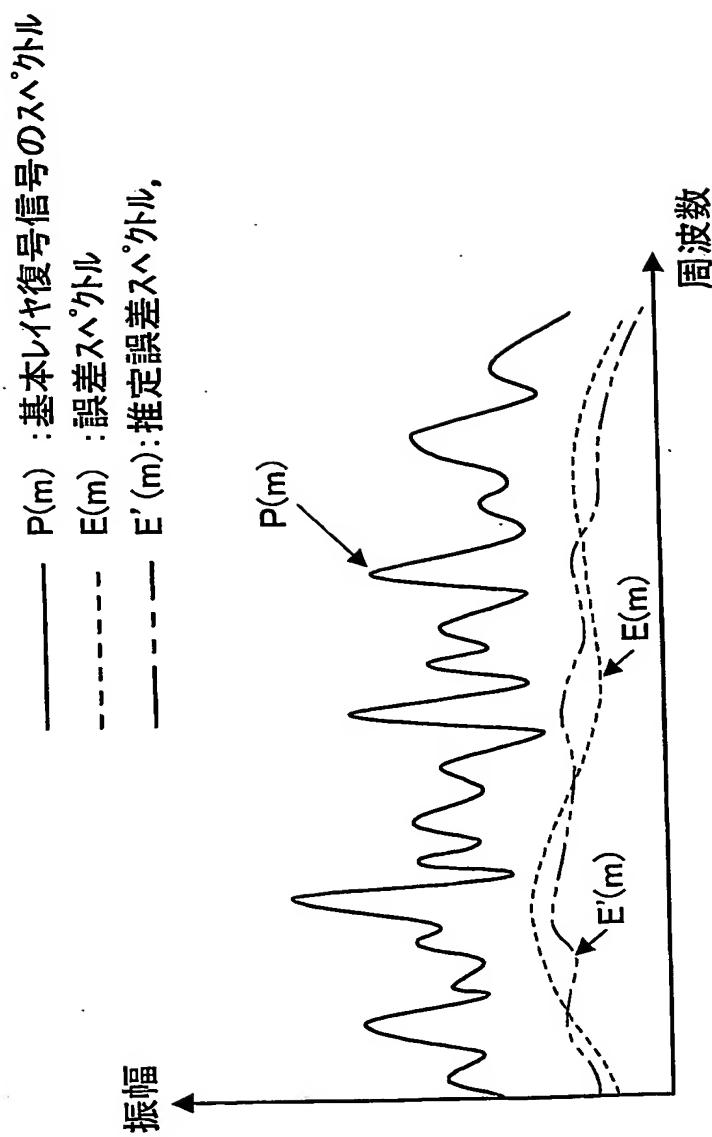


図30

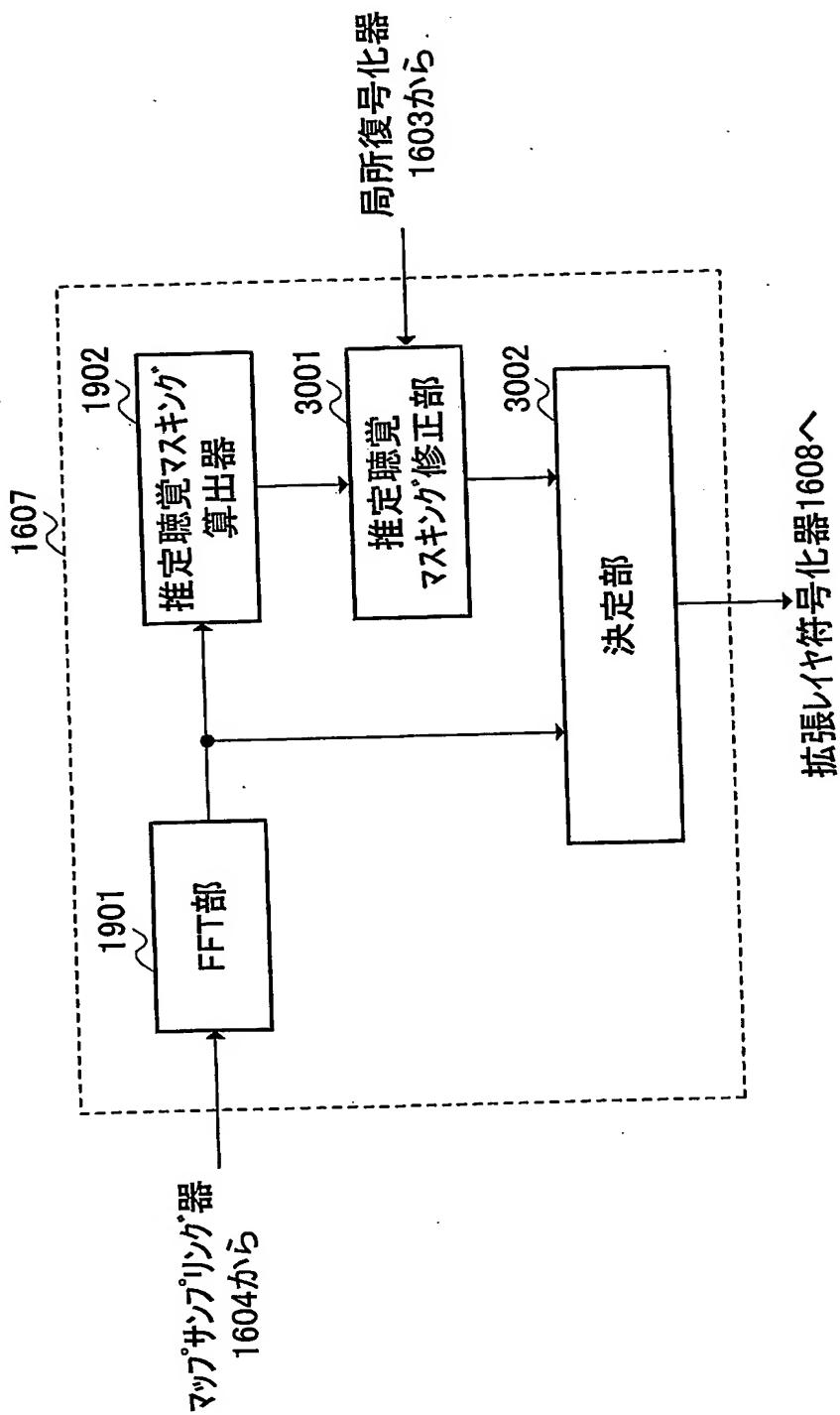


図31

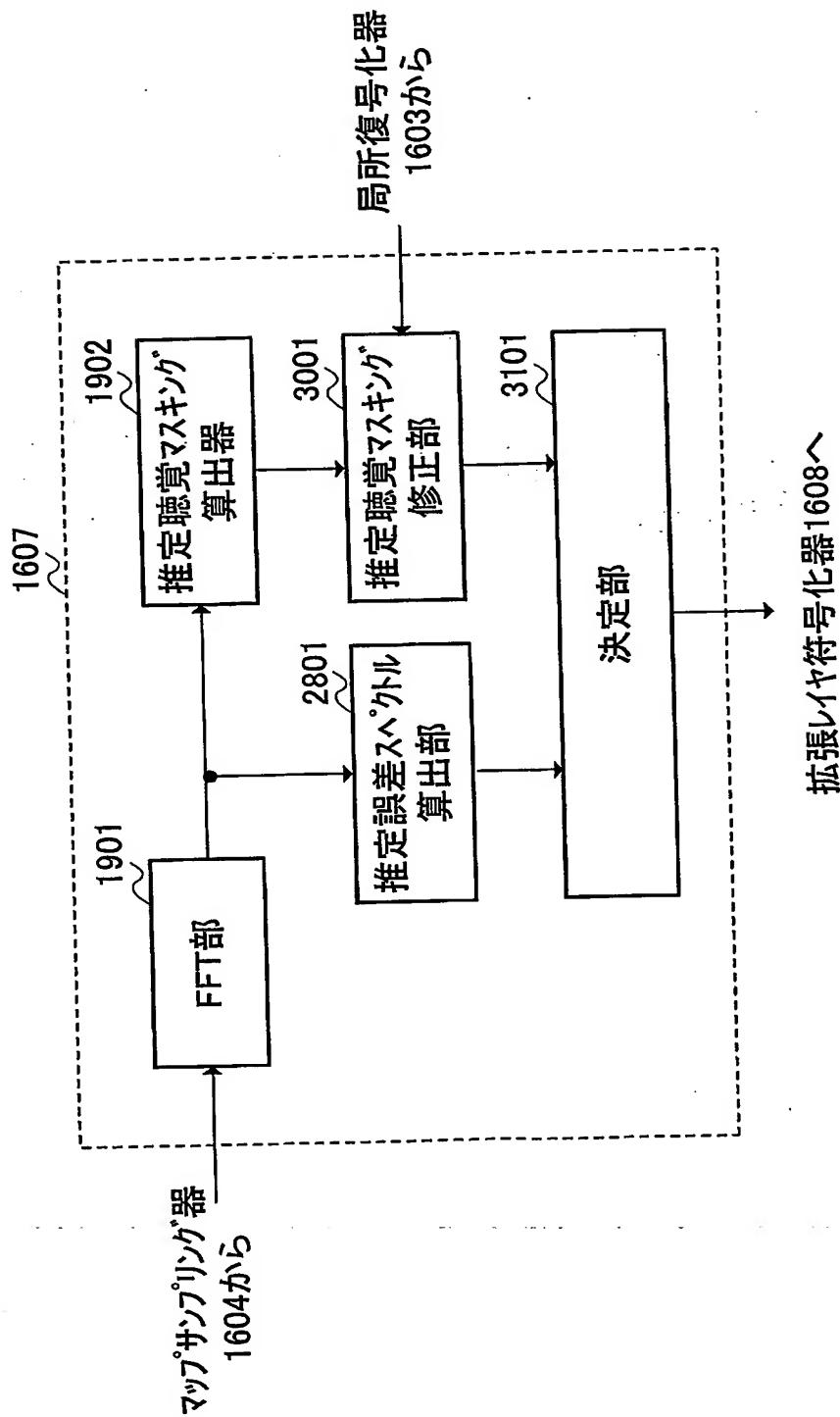


図32

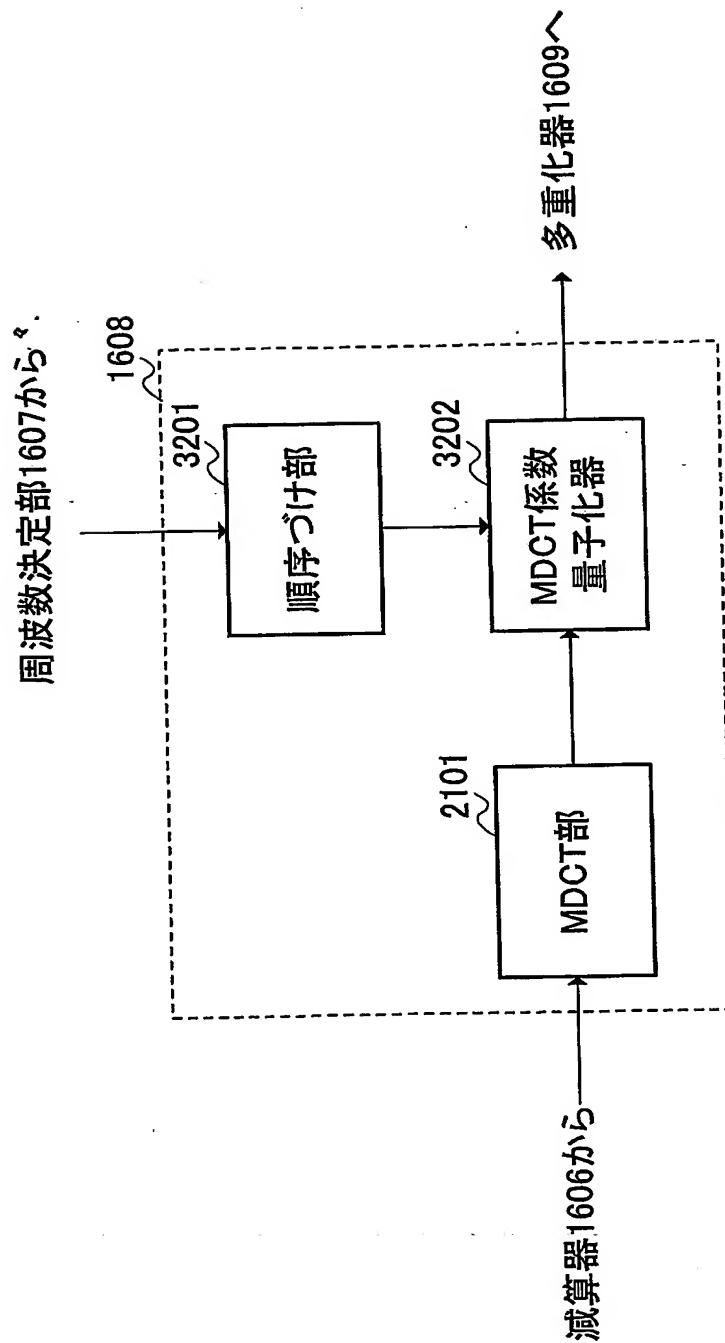


図33

32/39

| 周波数<br>(m) | 推定歪値<br>D(m) | 順序 |
|------------|--------------|----|
| 1          | 24.0         | 5  |
| 3          | 16.8         | 7  |
| 4          | 35.9         | 3  |
| 7          | 147.2        | 1  |
| 8          | 135.8        | 2  |
| 9          | 26.9         | 4  |
| 11         | 23.0         | 6  |
| 12         | 12.9         | 8  |

図34

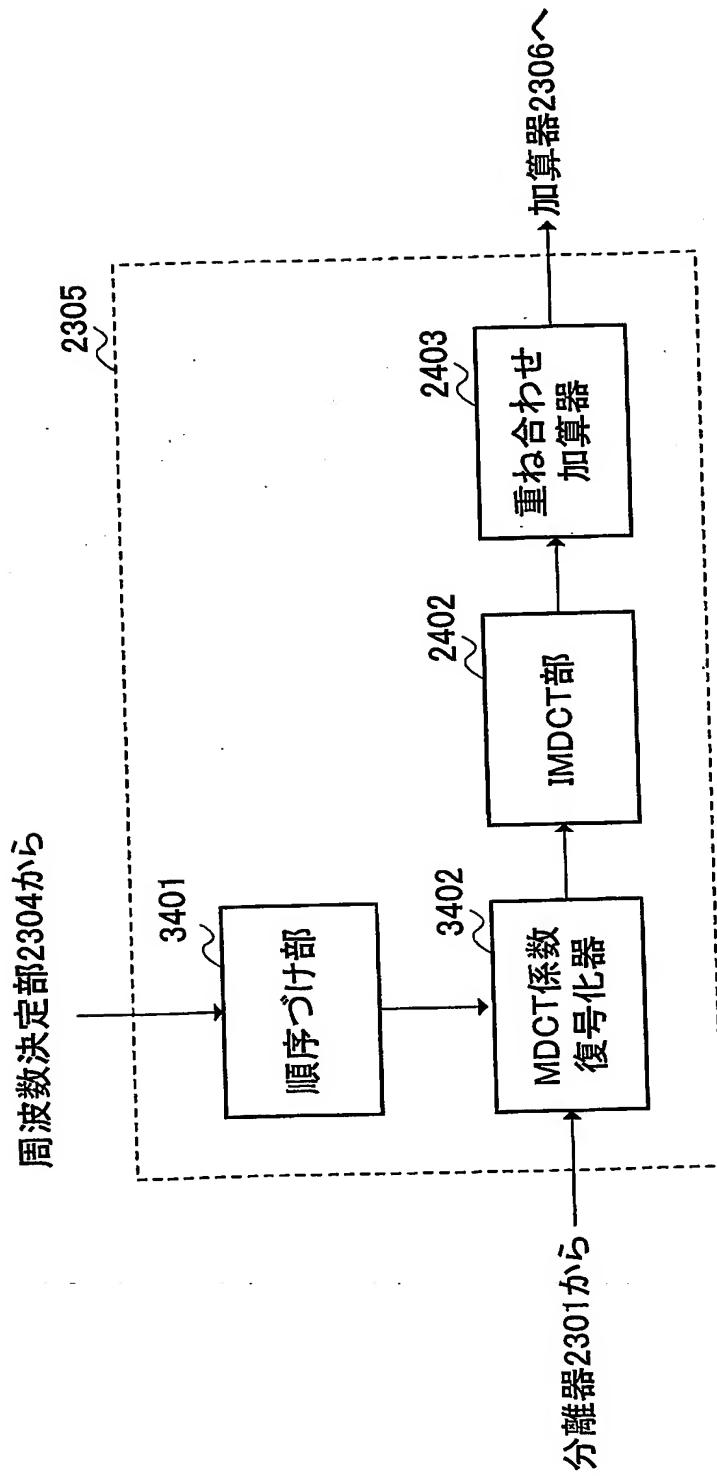


図35

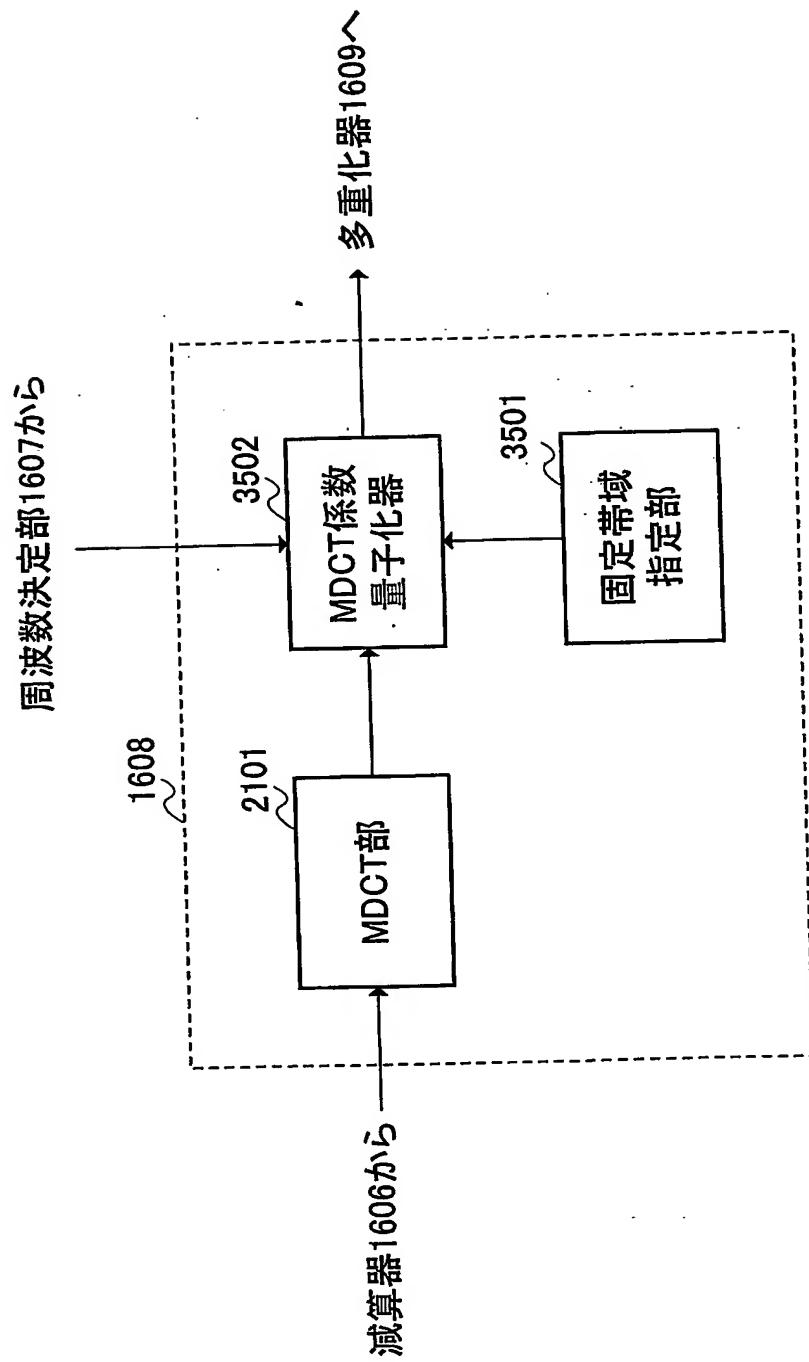


図36

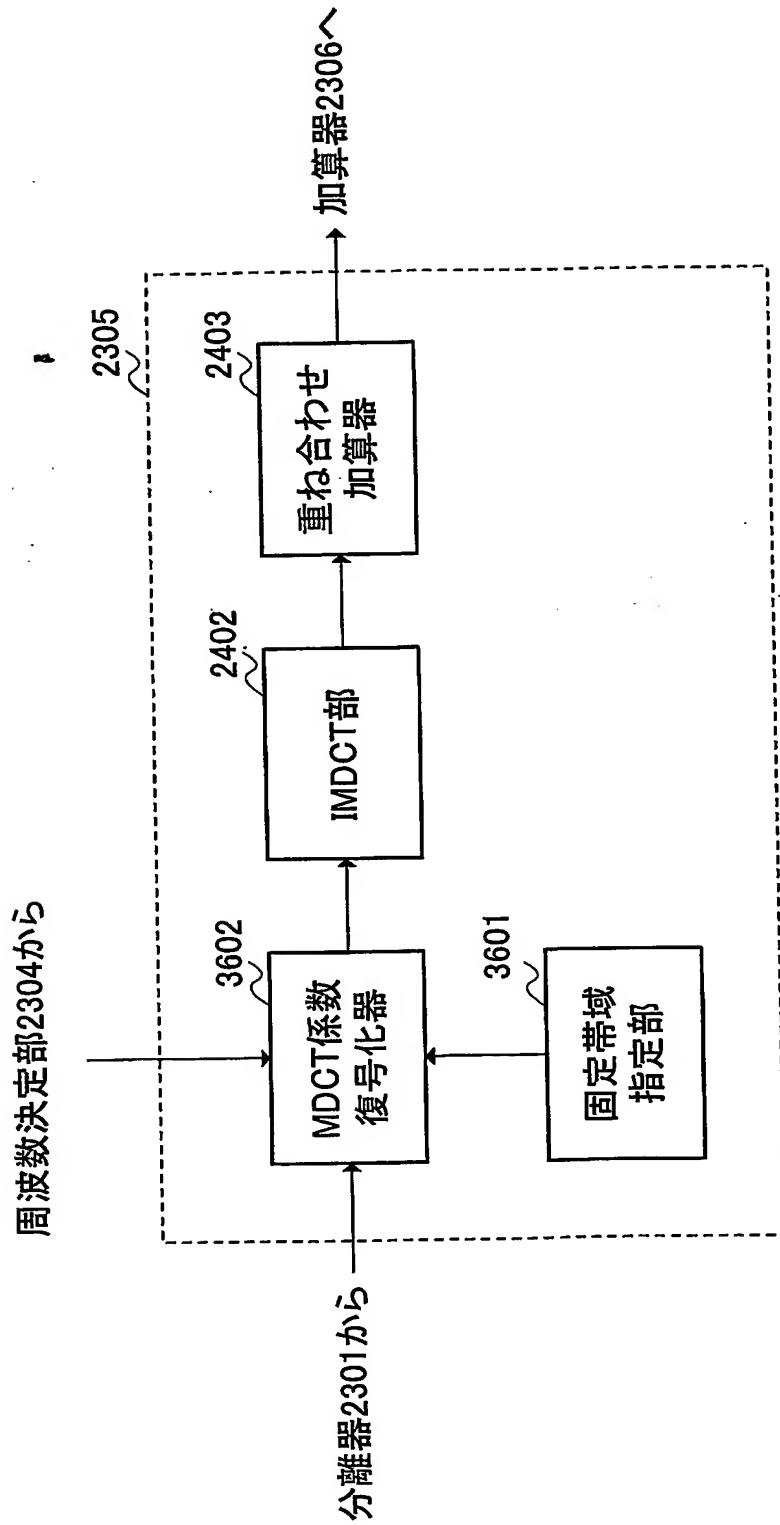
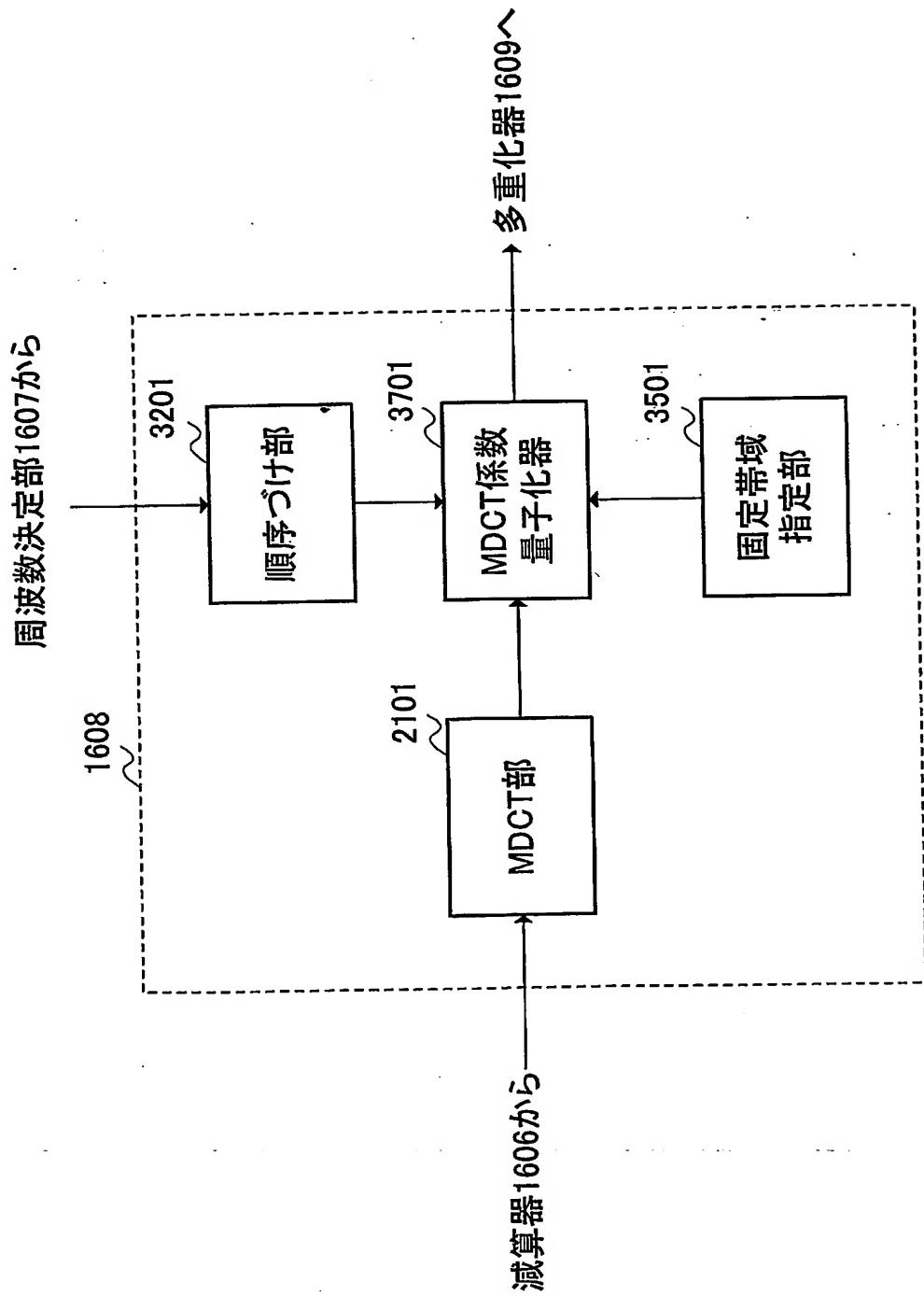


図37



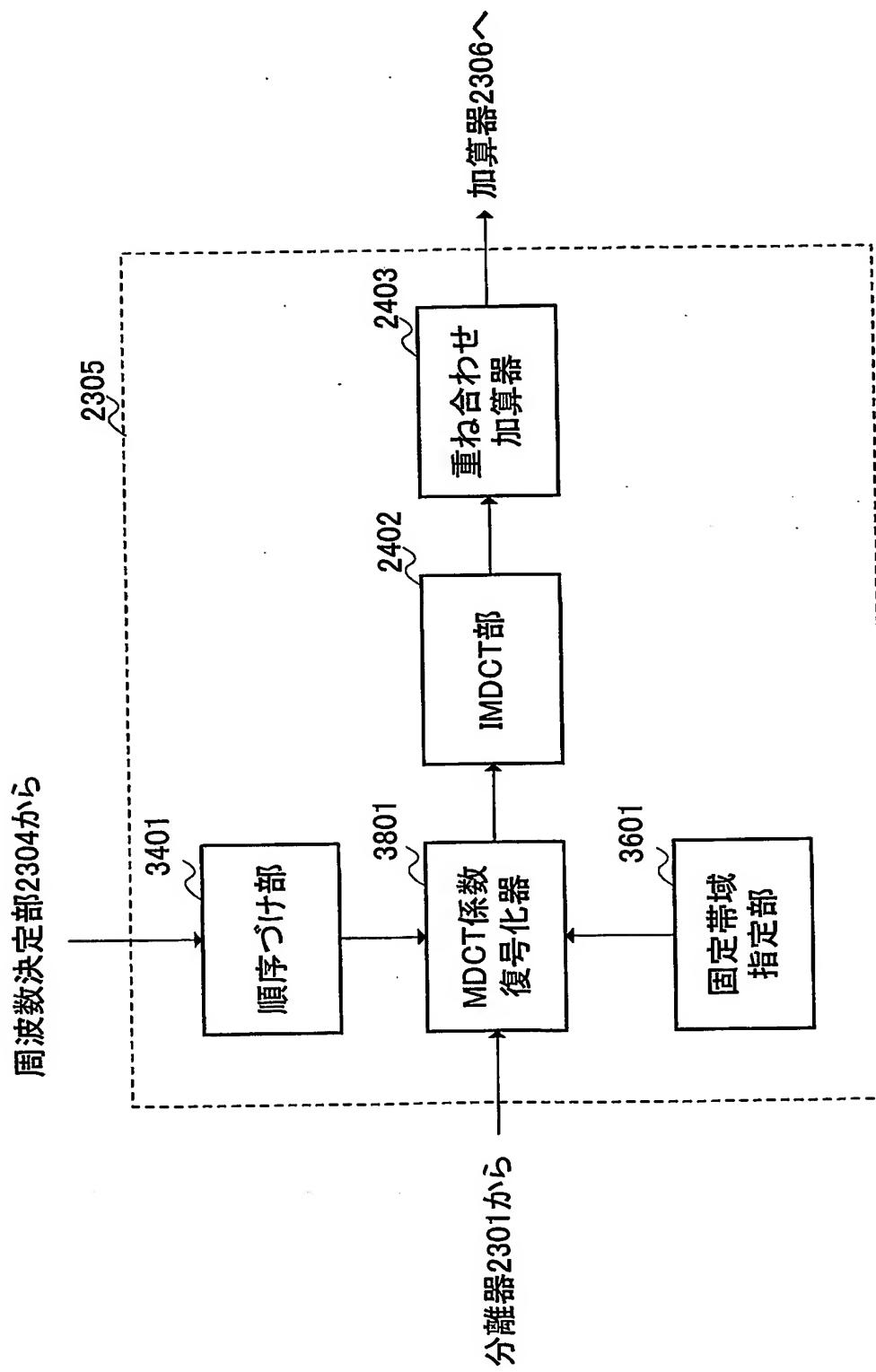


図39

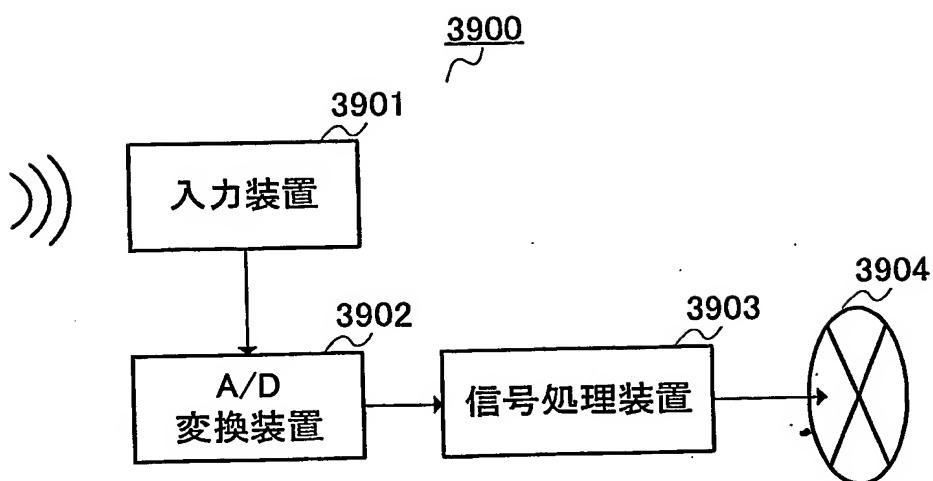


図40

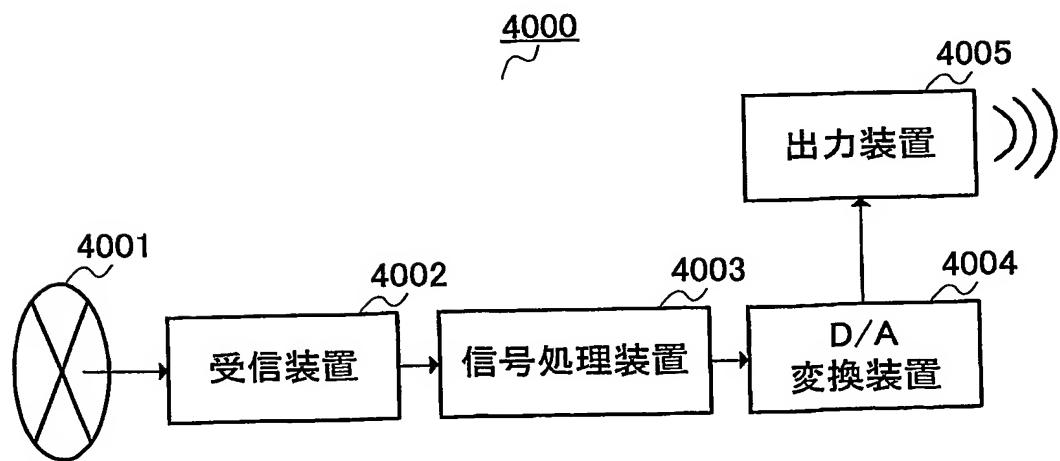


図41

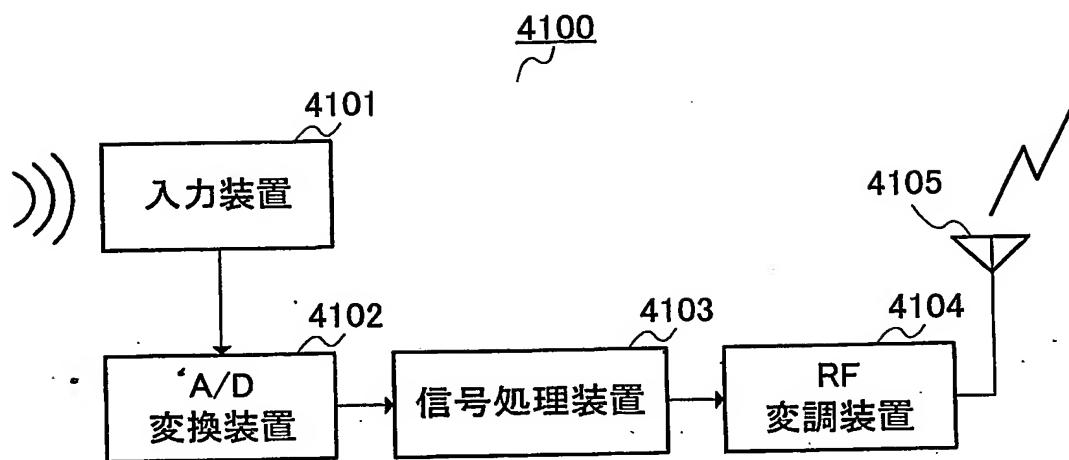


図42

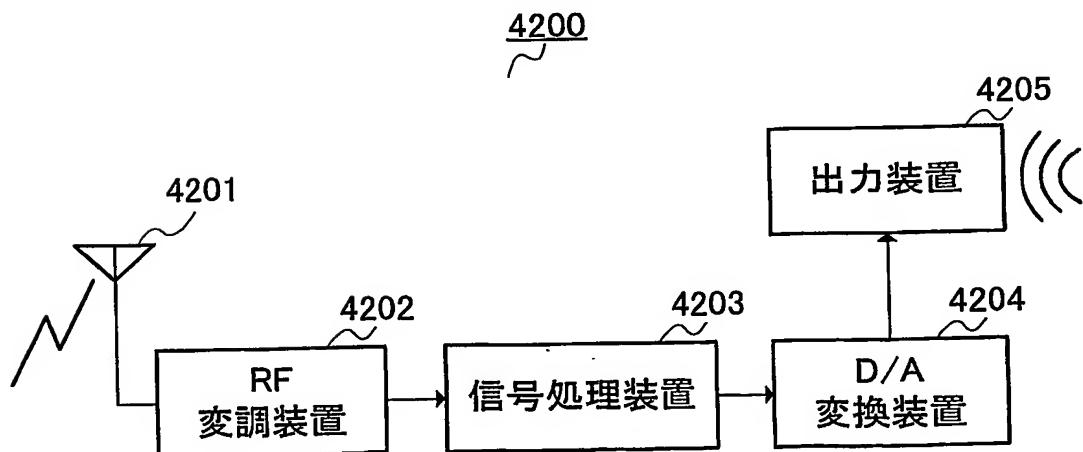


図43

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP03/05419

## A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl<sup>7</sup> G10L19/00

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl<sup>7</sup> G10L19/00

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

|                           |           |                            |           |
|---------------------------|-----------|----------------------------|-----------|
| Jitsuyo Shinan Koho       | 1922-1996 | Toroku Jitsuyo Shinan Koho | 1994-2003 |
| Kokai Jitsuyo Shinan Koho | 1971-2003 | Jitsuyo Shinan Toroku Koho | 1996-2003 |

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages             | Relevant to claim No.                 |
|-----------|--|---------------------------------------|
| X         | JP 11-30997 A (NEC Corp.),<br>02 February, 1999 (02.02.99),<br>& US 6208957 B1 & CA 2242437 A1 | 1, 2, 5-9, 18,<br>19, 22-26,<br>35-42 |
| Y         | & EP 890943 A2   | 3, 4, 10-17,<br>20, 21, 27-34         |
| Y         | JP 8-46517 A (Sony Corp.),<br>16 February, 1996 (16.02.96),<br>(Family: none)                  | 3, 4, 10-17,<br>20, 21, 27-34         |
| Y         | JP 11-251917 A (Sony Corp.),<br>17 September, 1999 (17.09.99),<br>(Family: none)               | 3, 4, 10-17,<br>20, 21, 27-34         |

 Further documents are listed in the continuation of Box C. See patent family annex.

\* Special categories of cited documents:  
 "A" document defining the general state of the art which is not considered to be of particular relevance  
 "E" earlier document but published on or after the international filing date  
 "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)  
 "O" document referring to an oral disclosure, use, exhibition or other means  
 "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention  
 "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone  
 "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art  
 "&" document member of the same patent family

Date of the actual completion of the international search  
14 May, 2003 (14.05.03)Date of mailing of the international search report  
27 May, 2003 (27.05.03)Name and mailing address of the ISA/  
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

## A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int. Cl' G10L19/00

## B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (IPC))

Int. Cl' G10L19/00

## 最小限資料以外の資料で調査を行った分野に含まれるもの

|             |            |
|-------------|------------|
| 日本国実用新案公報   | 1922-1996年 |
| 日本国公開実用新案公報 | 1971-2003年 |
| 日本国登録実用新案公報 | 1994-2003年 |
| 日本国実用新案登録公報 | 1996-2003年 |

## 国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

## C. 関連すると認められる文献

| 引用文献の<br>カテゴリー* | 引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示  | 関連する<br>請求の範囲の番号                      |
|-----------------|--|---------------------------------------|
| X               | JP 11-30997 A (日本電気株式会社), 1999.02.02<br>& US 6208957 B1 & CA 2242437 A1 & EP 890943 A2 | 1, 2, 5-9, 18,<br>19, 22-26,<br>35-42 |
| Y               |  | 3, 4, 10-17,<br>20, 21, 27-34         |
| Y               | JP 8-46517 A (ソニー株式会社), 1996.02.16<br>(ファミリーなし)  | 3, 4, 10-17,<br>20, 21, 27-34         |
| Y               | JP 11-251917 A (ソニー株式会社), 1999.09.17<br>(ファミリーなし)                                      | 3, 4, 10-17,<br>20, 21, 27-34         |

 C欄の続きにも文献が列挙されている。 パテントファミリーに関する別紙を参照。

## \* 引用文献のカテゴリー

「A」特に関連のある文献ではなく、一般的技術水準を示すもの  
 「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの  
 「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)  
 「O」口頭による開示、使用、展示等に言及する文献  
 「P」国際出願日前で、かつ優先権の主張の基礎となる出願

## の日の後に公表された文献

「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの  
 「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの  
 「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの  
 「&」同一パテントファミリー文献

## 国際調査を完了した日

14. 05. 03

## 国際調査報告の発送日

27.05.03

## 国際調査機関の名称及びあて先

日本国特許庁 (ISA/JP)  
 郵便番号 100-8915  
 東京都千代田区霞が関三丁目4番3号

## 特許庁審査官 (権限のある職員)

渡邊 聰



5 C

8622

電話番号 03-3581-1101 内線 3540